# THE PROBLEM OF ESTIMATION

# THE PROBLEM OF ESTIMATION

A SEVENTEENTH - CENTURY CONTRO-
VERSY AND ITS BEARING ON MODERN
STATISTICAL QUESTIONS, ESPECIALLY
INDEX-NUMBERS

BY

## CORREA MOYLAN WALSH

AUTHOR OF
"THE MEASUREMENT OF GENERAL EXCHANGE-VALUE."

# PREFACE

This little work has been written in the interest of harmony. In mathematics disputes must soon come to an end, when the one side is proved and the other disproved. And where mathematics enter into economics, it would seem that little room could be left for long-continued disputation. It is therefore somewhat surprising that one economist after another takes up the subject of index-numbers, potters over it for a while, differs from the rest if he can, and then drops it. And so nearly sixty years have gone by since Jevons first brought mathematics to bear upon this question, and still economists are at loggerheads over it. Yet index-numbers involve the use of means and averages, and these being a purely mathematical element, demonstration ought soon to be reached, and then agreement should speedily follow.

# CONTENTS

# THE

# PROBLEM OF ESTIMATION

*(A Seventeenth-Century Controversy and its bearing on Modern Statistical Questions, especially Index-Numbers.)*

## I.—THE CONTROVERSY.

AMONG the problems which it was customary for scientists to set to one another two and three centuries ago, an interesting one was started at Florence in 1627. A participant was Galileo, and in his collected works are usually included not only his own letters on the subject, but those of two others, a certain Nozzolini and the mathematician Castelli.[1]

The question propounded was this: If a horse worth 100 crowns is estimated by one person at 1,000 and by another at ten, which of these two estimates is the less erroneous, or are they equally erroneous? The problem may be varied thus: If an article worth 100 crowns is estimated by one person at 1,000, what lower estimate would be equally erroneous; or if the first estimate were ten, what higher estimate would be equally erroneous? It admits also of a further variation: one person estimates an article at 1,000 crowns, and another estimates it at ten; what must be its value in order that these errors be equal? Here it is evident that the answer is some sort of a mean between 1,000 and ten. In the second form of the problem the mean is given as 100, and the question concerns the extremes. In the original form the question is what relation the proper mean between 1,000 and ten has to the true value 100. Thus the problem of estimation is a question of means, and its solution involves the finding of the kind of mean suitable for equalising errors above and below the true quantity.

---

[1] References here are to the *Opere di Galileo Galilei* published at Milan in 1832. The dispute occupies twenty finely printed, double-columned, large octavo pages: Vol. I. pp. 57–77.

Evidently, we may say at once, the harmonic mean is ruled out of court ; for if one person estimates the article at half or at less than half its true value, the other person could not equal his error, according to this mean, except by estimating it at infinity or at some imaginary negative value, which has no meaning in the matter before us.   For the harmonic mean $m$ between two quantities $a$ and $b$ is thus algebraically expressed, $m = 2ab/(a + b)$ ; and now, if the mean and one of the extremes be given, the other extreme is thus expressed, $a = mb/(2 b - m)$, wherefore, when $b$ is the lower extreme, if it is half the mean, we have $a = mb/o = \infty$, and if it is less than half the mean, then $2b - m$ is a negative quantity, and the quotient is a negative quantity.   But a negative quantity, no matter what be its absolute size, cannot be greater than a positive quantity ; and yet it is offered as the greater extreme ; which is absurd.

The seventeenth-century disputants were well acquainted with the three classic means, the arithmetic, the geometric, and the harmonic, not to mention certain others.   Yet not a word do they say about the harmonic mean : their dispute narrowed itself between the arithmetic and the geometric. Galileo and Castelli advocated the former, Nozzolini alone stood out for the latter.   Perhaps it was the unwieldiness of the harmonic mean, or the less frequency of its use in general, that moved them to ignore it.   But perhaps, at least with some of them, a consciousness of its unsuitableness, as just pointed out, was a contributory cause of its exclusion.

Galileo, in a later stage of the dispute, tells us that when the question was first presented to him, he at once jumped to the answer, that the estimate of 1,000 crowns contained a much greater error than that of ten, since the gain in that case was 900 and the loss in the other only ninety.   That is, he was inclined, as most people are, to use the arithmetic mean.   But when he was informed that men of intelligence had raised a controversy over this question, he devoted more attention to it, and soon changed his mind, perceiving that there was more to be said in favour of the geometric mean. In his first letter he made a very simple argument.   The example offered might seem extravagant enough ; but he considered one still more extravagant.   He supposed that one estimate of an article worth 100 crowns was of 200 crowns, and the other was of only one crown.   If he employed the arithmetic mean, he would have to regard the former estimate as less erroneous than the latter.   Yet the former estimate was only twice as great as the true value, involving an error

not infrequently made, while the latter estimate was 100 times less than the true value, committing an error altogether unreasonable. Here, we may remark, Galileo abandoned difference for proportion. In the new supposition, the difference (and, we may say even, the distance) of the first estimate from the true value was a trifle greater than that of the second, but the ratio of the first estimate to the true value was much less than the ratio of the true value to the second. The problem, he concluded, is one to be solved by proportion ; and therefore the geometric mean is the right one for the case. And the original question, which set the one estimate as ten times too great and the other as ten times too little, he answered by declaring these estimates equally erroneous.

Nozzolini, unlike Galileo, stuck to his first inclination to use the arithmetic mean. The estimate of 1,000 crowns he unreservedly pronounced ten times more erroneous than the estimate of ten, since the former departs from the true value by 900 and the latter by only ninety crowns ; and to make their errors equal, he maintained the horse must be worth 505 crowns. But knowing of the geometric mean, which was suggested by the statement of the problem itself, and suspecting that some might advocate it, he proceeded to give an argument for the arithmetic mean. He put it on the ground that this is a case for commutative justice and not for distributive justice, and it has been established since the time of Aristotle that the geometric mean is the proper one to use in distributive justice, but in commutative justice (defined as a mean between gain and loss [2]) the proper mean is the arithmetic. But his reason for assigning the affair to commutative justice is by no means clear. He maintained that, as a purchase is a ratified estimate, so an estimate is nothing but an unratified purchase ; and he observed that in purchases and sales all errors are by merchants reckoned arithmetically. Yet he mentions many transactions in which merchants make use of the rule of three (which is a geometric process), and especially he notices that gains and losses are divided by partners in proportion to the capital they invested, wherein the geometric method is used. These, however, he says, are cases of distribution, and there-fore they rightly employ the geometric mean. Still it remains that, although a question of estimation may not be a subject for distributive justice, it is not shown that it is a

---

[2] *Nic. Ethics*, V. iv., rather between pelf and penalty.

matter of commutation, or balancing of gains and losses ; nor is it proved that the geometric mean is confined to questions of distribution.

When Nozzolini wrote this first letter, he was unaware of Galileo's opinion.   After Galileo's letter was sent him, he returned to the charge.   Galileo's still more extravagant case did not bother him in the least.   The man who estimates a hundred-crown article at one crown commits, he says, a smaller error than the man who estimates it at 200, as 99 is less than 100.   The argument about proportion he rejects on the ground that it changes the measure.   In all comparisons of magnitude, he asserts, we must employ only one measure.   We cannot say that two distances are equal because the one is ten yards and the other ten rods.   And now when Galileo adduces the fact that in the original example the one estimate is ten times too great and the other ten times too little and therefore holds them to be equal, Nozzolini retorts that he has changed the measure, since the first estimate is distant from the truth by ten large measures of 100 crowns and the second estimate is distant from the truth by ten small measures of ten crowns ; wherefore he likens it to saying that at Florence the campanile is the same distance from the cathedral as is the church of San Giovanni because the latter's distance is ten steps of a giant and the former's is ten steps of a baby.   The argument, of course, goes too far ; for by it a considerable amount of our ordinary use of proportion, or the equationing of ratios, would be quashed.   Nozzolini has continued to speak of distances, or differences, when the question has been put in the form of ratios.   In proportions the measures are eliminated, and therefore, as Galileo will show us in his reply, various measures can be used when we compare the ratios of the errors of estimation committed concerning them.

In conclusion, to bring some hilarity into the too serious subject, Nozzolini appealed to a forum which he declared to be competent and which every day adjudicates such questions, and never hesitates in its decision.   He had noticed, he wrote, that butchers often laid wagers among themselves and with countrymen about the weight of pigs or calves, and they decided the question always in favour of the one who came nearest to the true weight found when they weighed the animal.   We have a similar proceeding in our country fairs, where often a charge is made for guesses over the weight of some huge hog or pumpkin, and the one who guesses nearest to the truth gets the prize.   Here, apparently,

appeal is made to common sense in favour of the arithmetic mean.

Castelli now entered the lists, having seen the preceding letters. He backed up Galileo's position with two very pithy variations on the new example. First, if an article worth 100 crowns is estimated by one person at 200, he asked Nozzolini what should be the estimate below 100 with an equal error. The arithmetic answer must be that the other estimate should be nothing, since the difference between zero and 100 is the same as between 100 and 200. But is it not absurd to adjudge the former error as great as the latter, especially as, if some one estimated the article at 300, and it was asked what the lower estimate must be to match this error, only a very ridiculous answer is indicated ? This, we may remark, is analogous to the argument whereby we at the commencement rejected the harmonic mean. It shows that the arithmetic mean involves the same absurdity from the other side. If one estimate is half or less than half the true value, the harmonic mean can assign no valid higher estimate with equal error. If one estimate is double or more than double the true value, the arithmetic mean can assign no valid lower estimate with equal error. The arithmetic mean is in this respect as faulty as the harmonic. Secondly, taking up Nozzolini's idea of the kinship between estimating and buying-and-selling, and supposing a hundred-crown article to be estimated at 199 crowns and bought for that sum by one party, and another equally valuable article to be estimated at and bought for one crown by another party (with equal error, according to Nozzolini), Castelli pointed to great inequality, in that, as he alleged, in the first transaction the gain and loss are 99 per cent. and in the second they are 9,900 per cent. Here he made a mistake through too great conciseness. The seller's gain in the first transaction was 99 per cent., but the buyer's loss was less than 50 per cent. ; and in the second the seller's loss was again 99 per cent., and the buyer's gain was 9,900 per cent. Still, though the seller's gain and loss in each case have the same percentage, the loss and gain of the two buyers have entirely different percentages. Compared with the amounts they ventured, their gain and loss are utterly unequal.

Nozzolini answered Castelli in another letter. We may consider first his reply to the second objection, and carry it to its end. Here, he says, Castelli has unwarrantably altered the terms of the proposition. For at first in reckoning the gain of him who sold the hundred-crown article for 199 crowns

at 99 per cent., he went from the just price to the unjust price ; and then in reckoning the gain of the purchaser of the similar article for one crown at 9,900 per cent., he went contrariwise from the unjust price to the just price.   Nozzolini contended that we should go in both cases from the just price to the unjust price ; and therefore he relied on the fact that in each of the transactions the seller lost and gained the same percentage, wherein equality is manifested according to the arithmetic mean.   Rather, we might claim, it was Nozzolini who altered the propositions by going in opposite directions ; for he went up from the mean to the higher extreme in the one case, and in the other he went down from the mean to the lower extreme ; whereas Castelli went up from the lower extreme to the mean and continued going up from the mean to the higher extreme, although he happened to perform the latter ascent first.[3]

We have no rejoinder from Castelli, but Galileo takes his place.   While not admitting that gains and losses in actual transactions are the proper measure of error in the estimates upon which they are founded, he replied that the advantage of gaining 99 per cent. on one's capital is by no means equal to the disadvantage of losing 99 per cent. of one's capital. Absolute gains and losses are not proper measures of business capacity, in which the ability to make accurate estimates plays a considerable part ; or else we should have to say that a merchant who increased his capital from 1,000 crowns to 2,000 was a better man of business than one who in the same time increased a capital of 100 crowns to 1,000 ; which is not so, as the former has increased his capital only 100 per cent., while the latter has increased his 900 per cent.   To prove that the incapacity of him who reduces his capital 99 per cent., from 100 to one, is much greater than the capacity of him who raises his capital 100 per cent., from 100 to 200, Galileo used this bit of demonstrative reasoning.   Evidently the former's incapacity is much greater than the incapacity of one who reduces his capital from two to one ; but this one's incapacity is exactly equal to the capacity of another who raises his capital from one to two ; and this other person's

[3] Nozzolini added another counter-argument.   He said that Castelli's problem was stated in the arithmetic progression 1, 100, 199, and therefore it was impossible to get a geometric progression out of it (p. 62B).   What was meant by this, it is difficult to see. As well might it be said that the original problem was stated in the geometric progression 10, 100, 1,000, and therefore the arithmetic progression could not be got out of it.

capacity is equal to the capacity of the person who raises
his capital from 100 to 200 ; therefore the first one's incapa-
city, being much greater than a capacity equal to this last
one's capacity, is much greater than this last one's capacity
(p. 69A below).    The argument hinges on the equality of the
opposite qualities of two persons, of whom the one decreases
his capital from two to one and the other increases his capital
from one to two.    This is evident.    In practice, however, we
reckon these movements from their starting points, and so go
in opposite directions, saying the one has lost 50 per cent.
and the other has gained 100 per cent. ; wherein there is no
appearance of equality.    But the loser here is a person who
was 100 per cent. better off than he now is, and the gainer is
a person who now is 100 per cent. better off than he was ; or,
to turn both movements around, the loser is now 50 per cent.
worse off than he was, and the gainer was 50 per cent. worse
off than he now is.    The appearance of equality is restored
by reckoning always in the same direction, in each case from
a higher to a lower position, or from a lower to a higher, as is
done in the geometric mode of reckoning.    And so in esti-
mates, the geometric mean, indicating the true value, stands
in a proportion $x : m :: m : y$, in which $x$ and $y$ represent
estimates geometrically equally above and below the true
position ;   whence $\frac{x}{m} = \frac{m}{y}$ or $\frac{m}{x} = \frac{y}{m}$. The  mean  $m$  must
always be either in the numerator or the denominator, and
cannot be in one position throughout in the same proportion.
Thus an estimate above the truth may be represented by a
ratio of the estimate to the mean, and then the estimate
below the truth must be represented by a ratio of the mean
to the estimate ; or the former may be represented by a ratio
of the mean to the estimate, and then the latter must be
represented by a ratio of the estimate to the mean.    There-
fore, in geometrically comparing errors on opposite sides of
the truth, we must always go in the same direction, whichever
be the direction first chosen, as Castelli and Galileo did, and
not in opposite directions, as Nozzolini would have had
them do.
    To Castelli's first objection, which, when an estimate is
100 per cent. too high, required Nozzolini to place the
equally wrong estimate on the other side 100 per cent.
below the true value at zero, Nozzolini replied that he would
do nothing of the sort.    He would still do what he had done
in boyhood when he had occasion to meet excessive estimates
with the corresponding estimates below the correct ones.

For when in the market he had been asked two farthings for a pear which he knew was selling for one farthing, he had been in the habit of countering, not by offering nothing for one pear, which he knew would be absurd, but by demanding two pears for one farthing. And so now he maintains that when the owner of the hundred-crown article estimates it at 200, virtually demanding two payments for one article, the corresponding lower estimate of the purchaser should be to demand two such articles for 100 crowns, thereby virtually offering one payment for two articles. Here, not in only the one, but in both the cases, he adds, is hidden the zero complained of. For, as the over-estimating seller wants two payments and offers the article for the one and nothing for the other, only so the under-estimating purchaser wants two articles and offers payment for the one and nothing for the other (p. 62A).

Here Nozzolini seems to give away his case, since to demand two articles for 100 crowns is the same as to demand each of them for fifty, so that the double estimate is met by a half estimate, as required by the geometric mean. But at the end of the letter, almost as an afterthought, and again incidentally in his reply to Galileo's repetition of the objection, he discloses his meaning. He has converted the question from one of estimation to one of gain and loss in a business transaction, and then very lightly jumped back to the question of estimating. Thus, reverting to his idea that an estimate is nothing but an unratified purchase, and letting go of estimation, and considering only what happens in buying and selling, he says that when a merchant sells an article for 20 per cent. more than it is worth, the next time he sells a similar article to the same purchaser, if he wishes to make up for his first injustice, and to equalise his price above by a price below the true value, he should now ask 20 per cent. less than it is worth—that is, if he had once sold a hundred-crown article for 120 crowns, he should now sell it for eighty ; which is as required by the arithmetic mean. And he points out that it would be absurd for the unjust seller of the hundred-crown horse at 1,000 crowns to think that he could equalise the matter and make up for his injustice by selling to the same purchaser another hundred-crown horse for only ten crowns, in accordance with the geometric mean. No, he says, in this case the unjust seller, to ease his conscience, must sell ten such horses for ten crowns apiece (p. 63). This is perfectly evident of purchases and sales ; and now Nozzolini claims

that the same must be true of estimates. He demands an estimate of ten horses at ten crowns apiece in order to have the same total amount of error in these estimates below the true value as was contained in the single estimate ten times above the true value, which cannot be equalled in a single estimate below the truth, even though, he says, the horse were estimated as worth less than a grain of sand (pp. 64A, 73B).

Thus Nozzolini admits that the error in an estimate twice, or more than twice, the true value cannot be equalled by the error of a single estimate below the true value ; but he thinks he gets rid of the difficulty by distributing the error over as many estimates as the first estimate was times greater than the true value. We may remark that in this way the harmonic mean might be rehabilitated, since an estimate less than half the true value could be met on the upper side by distributing the error over two or more positive, finite, and not very large estimates. In general, against both these means, we may argue that, although there is some resemblance of an estimate to a purchase, there is also some difference, and here the difference tells. For, if we allow a single hundred-crown horse to be estimated along with nine other imaginary horses supposed to be its duplicates, when the first estimator, thinking of an actual horse, estimates it ten times too high, and the second estimator, in order to equal his error on the lower side, estimates the actual horse and the nine imaginary ones at ten crowns each, thereby in fact committing a total error of 900 crowns, equal to the other's, there is no reason why the first estimator should not now do the same, and so when he estimates the nine additional imaginary horses at 1,000 crowns each, the error of his estimate still is, in this way of computing error, ten times greater than that of the other's. At this rate the second party will never be able to overtake the first in error ; for however many imaginary horses he invokes to estimate below the true value, the same can the first party employ to estimate above the true value.

Galileo, for his part, after first making a reply which did not altogether reach the mark, as he seems to have failed quite to grasp Nozzolini's argument, made another reply, which left Nozzolini with nothing to say, for he first ignored it, rebutting only the less successful reply, and afterwards refused to consider it. Galileo claims that the problem of estimation need not be confined to values, but may be made of other quantities in which the question cannot be trans-

ferred to one of gain and loss in an exchange. Thus, altering Castelli's example, he supposes a tower 100 feet high to be estimated 200 feet high, and asks what estimate below the true height will be equal in error to that. Certainly it cannot be an estimate of zero feet ; for that would raze the tower to the ground. And if Nozzolini says the other estimate must be of two towers each 100 feet high so made as to get the total error of 100 feet, he must say either that one of the towers is 100 feet high and the other of no height at all (and so no tower), or that they are each of them 50 feet high (p. 68B). In the latter case, as the imaginary tower does not concern us, we may accept the estimate of the tower in question, which is as required by the geometric mean. Or, if Nozzolini insists on retaining the additional tower, we may say that he has altered the problem, which was about a single tower. And we may repeat, if the one estimate can deal with two towers, why cannot the other ?

What has in the above two instances been quoted of Galileo, was made in a postscript to his only other letter. In the letter itself, written before he had information of these, he replied only to Nozzolini's earlier letters. After commenting on Nozzolini's reference to commutative and distributive justice, which had no weight with him, and repeating his own argument from extravagant instances by another extravagant supposition (of a bagful of coins being estimated double and very little), and raising Castelli's first objection in a slightly different form in ignorance of Castelli's having already made it, and denying that any gain and loss need be considered, as they do not enter when other quantities not values are estimated, he broaches a broader treatment of the subject. He points out that Nozzolini's fundamental error in measuring errors of estimation was in considering only the absolute magnitudes of the errors and in taking no account of the relations between the estimates and the things estimated. This procedure leads to all sorts of absurdities, among which is the possibility that most ignorant appraisers may have the preference given them over appraisers of the greatest experience and perspicacity. For if any one should estimate the value of a nut (ten of which are sold for a farthing) at a crown, he would by everybody be considered a most extravagant appraiser ; but if any one should make an error of a crown in estimating the value of a jewel worth four or five thousand, he would by everybody be considered a most accurate appraiser. Galileo

believed that Nozzolini would not judge otherwise ; and yet his reasoning would require him to consider the former the better appraiser. And to show that we cannot really judge errors of estimation without relation to the size of the things estimated, he supposes the case of one who in estimating the height of a mountain makes an error of 100 yards, and of another who in estimating the weight of a bullock makes an error of ten pounds. Here he points out that it is impossible to compare these errors, for the data supplied are insufficient. But if we add that the true height of the mountain is 1,000 yards, and the true weight of the bullock is 100 pounds, then it is possible to compare the errors ; but we cannot say the error of the former is the greater on account of 100 being greater than ten, as there is no comparison between 100 yards and ten pounds ; or if we judge simply by the absolute numbers, we might change the latter statement and say the appraiser of the bullock erred by 120 ounces, and then, as 120 is greater than 100, his error would, from the smaller, become the greater ; which is absurd. The correct answer is that the two errors are equal (provided, we must add, they were on the same side), because in each case the error is of 10 per cent.[4]

[4] Galileo did not make the proviso in the parenthesis, but it is understood in his argument. Of course Nozzolini would agree with him in the above. Nozzolini's real difference with Galileo lay in holding that an equal percentage of error above and below the true quantity makes an equal error. Sometimes, however, he placed the equality in the difference, or distance, from any true quantity. If a single tower 100 feet high is estimated at 120 feet and at 80 feet, he says the errors are equal for both these reasons, that the estimates each depart 20 feet from the true height, and that they each depart 20 per cent. from the true height. But these reasons are by no means the same, for, though here coinciding, they may diverge. This may be shown by repeating Galileo's back-and-forth argument, adapted to the present case. If a tower 100 feet high is estimated at 80 feet, and if another tower 80 feet high is estimated at 100 feet, the error in each of these estimations is of 20 feet, and Nozzolini must say they are equal. In fact, in the last paragraph of his last letter on the subject he expressly says that in the arithmetic method of measuring, nothing is considered but the distances (or differences), and that it is as much to go from the more to the less as from the less to the more, there being the same distance from 8 to 4 as from 4 to 8, just as it is the same distance from my house to yours as from yours to mine. According to this, the distance from 100 to 80 being the same as from 80 to 100, the

In measuring errors of estimation against Nozzolini's contrary opinion, Galileo insists that the unit of measurement is indifferent, as shown by the above comparison between estimates of a height and of a weight. Taking another instance, Galileo offers as self-evident the proposition that if one person estimates the height of a tower at 150 feet, which afterwards is measured and found to be 100 feet high, and if another estimates the weight of a calf at 150 pounds, which afterwards is put on the scales and found to weigh 100 pounds, their errors are equal, although the measures are different in kind. He concluded that the measurement of errors of estimation is generically different from the measurement of concrete things, because error is an abstraction. All measures, he maintains, must be of the same species with the things measured, as we measure times by a time, weights by a weight, values by a value. And so errors must be measured by an error, which is a ratio between the estimate and the true quantity, and not a concrete quantity itself. We cannot measure errors by so many pounds, feet, or crowns : we must measure them by the proportions of the pounds, feet, or crowns in the erroneous estimates to the pounds, feet, or crowns in the thing estimated. He therefore defines the measure of error in estimation as, abstractly, the general relation or habit which the false estimate has to the true quantity of the thing estimated (pp. 66, 67). Hence, he repeats, all reference to commutative and distributive justice is not to the point ; and he opines that, had the question originally been put about errors in estimating the height of a giant, supposed to be 10 feet, and one estimate was supposed to be 100 feet and another 1 foot, Nozzolini would never have argued as he did, and would have seen that those errors are equally exorbitant. But if Nozzolini persists in his opinion, Galileo warns him, he will have to maintain many ridiculous things—as that he makes a better estimate of a heap of 1,000 coins who says he thinks there may be two or three,

errors must be equal. But again, according to the usual way of reckoning percentages, while the former error is of 20 per cent., the latter is of 25 per cent. Therefore the same Nozzolini must say the errors are unequal—and this time he would be wrong. Into such inconsistency did his application of an arithmetic process to a case of geometric proportion lead him. (Reasoning similar to Galileo's in this matter was employed, without knowledge of Galileo's precedence, by the writer in *The Measurement of General Exchange-Value*, pp. 351–2, cf. 236–9, 250–3.)

than he who judges them 2,000 ; or worse yet, he must
hold that he who estimates Monte Morello to be 10,000 feet
high (over twice its height) is a worse judge of height even
than he who says it is a hole in the ground ; for Galileo
never thought of Nozzolini's scheme of bringing other
mountains into the question.   In fact, he says that for
Nozzolini to equalise, in the original problem, the error of
the estimate of the hundred-crown horse at 1,000, he must
suppose the horse to have a harness put on worth 810
crowns and then the horse and the harness together (worth
910 crowns) to be estimated at ten crowns.   This suggestion
is little more absurd than the method Nozzolini did adopt
to avoid the difficulty.

Lastly, Galileo refers to Nozzolini's appeal to the forum
of the butchers, and points out that the butchers used the
arithmetic mean through ignorance, as they did not perceive
that equal differences above and below the true value do
not spell equal errors ; a reason for which was that, being
skilled in such matters, they generally made but small
mistakes, and in small mistakes the difference between the
arithmetic and the geometric means is trifling (*e.g.*, between
110 and 90, while the arithmetic mean is 100, the geometric
mean is over 99) ; whereas, were they confronted with a
case where somebody estimates a calf of 100 pounds at
200 pounds and another estimates it at an ounce, they never
would have given the prize to the latter.   Herewith Galileo
left the subject.

But Nozzolini had the last word, writing two more letters,
one of them the longest in the lot.   Herein, after insisting
on Aristotle's authority, he traversed all of Galileo's criti-
cisms, mostly repeating only what he had already said.
But to the last criticism about the butchers he replied that
no matter how small the difference between the arithmetic
and the geometric means in small errors, the right mean
ought to be adopted (p. 74B).   It is difficult to see what he
was driving at by this ; for it could hardly be maintained
that the butchers must have adopted the right mean, and
as they adopted the arithmetic mean, therefore this is
the right one.   The real reason why the butchers, and other
people in similar situations, adopt the arithmetic mean,
and will continue to adopt it, is because it is the simplest
and most convenient.[5]   Simple methods of measuring are

[5] So, to vary the example, children will in all probability always
maintain that a lesson-mark of 80 out of a possible 100 is twice as
bad as one of 90.   Yet their error can easily be shown by con-

constantly used in practice, even by persons who know
that some other more complex method is more precise,
provided the former is precise enough for practical pur-
poses—and Galileo showed no sign of wishing to interfere
with the practice of the butchers.[6] Where precision is
absolutely required is in theory.  So it was Nozzolini, and
not the butchers, from whom we have a right to demand
precision.

Another rejoinder was a general one to Galileo's method
of arguing from extravagant cases.  Extravagant cases,
very conveniently for his own contention, he would throw
out altogether, notwithstanding that he had admitted con-
sideration of them at the beginning, before he perceived
the havoc they would make with his side of the question.
For justification he appealed to the procedure in dialectical
disputes ; for in them, he said, if one party denies that
snow is white or that fire is hot, he is regarded as a fool
with whom no argument can be profitably made.  And so,
he says, if any one should estimate a hundred-pound pig
as weighing an ounce, or should consider a mountain a
hole in the ground, his estimate should be rejected as that
of a fool who cannot be argued with.  He holds that from
extravagant cases we cannot judge of the nature of a thing,
citing in example that because a drop of water is round

sidering the case of eleven scholars, the first of whom receives the
mark of 90, the next a mark twice as bad, the third a mark three
times as bad, and so on.  The tenth, according to the children,
will receive the mark of zero.  What will be the mark of the
eleventh ?  Again, it is evident that 50 is half as good a mark as
100.  Then the arithmetician must say that zero is half as good
a mark as 50.  But obviously the mark half as good as 50 is 25.
This is in accordance with the geometric measurement.

[6] In their wagers the only case in which the arithmetic mean is
likely to give a different result from the geometric is where two
estimates are equally above and below the measured weight ;
for then, while the geometric mean would award the prize to the
one who made the higher estimate, the arithmetic mean would
divide it between them.  But here the superiority of the higher
estimate would in most cases be so inappreciably small, that it
would generally fall short of the probable error in the measure-
ment of the true weight ; wherefore even theory would sanction
the practice of using the arithmetic mean in such cases.  And if
the condition is first stated, as is usually done, as being that the
one who comes nearest to the measured weight shall be considered
the winner, there is no room for dispute about any other mean
but the arithmetic.

(a very small instance), we cannot infer that a barrelful of water has the nature of roundness (p. 72A). This is a rather unfortunate example, as we should do exactly what he denies : we should infer that a barrelful, or a whole oceanful, of water, if as free in space as is a drop of rain in the air, would assume the spherical shape, just as, and for the same reason that, the drop does. He admits that he had himself entertained the supposition of some such extravagant estimates, as of the hundred-crown horse being estimated at only ten crowns ; but now he says he had all along supposed the estimator had some reason for his estimate, as that he suspected some disease in the animal. A more humiliating admission is another which he now makes, to the effect that the person who should estimate a heap of 1,000 coins on the table before him at two or three makes a greater error than the one who estimates them at 2,000, although the former's estimate is nearer the truth. But he demands the privilege of throwing out all such foolish estimates, and of considering only the more judicious closer ones. He overlooks that Galileo did not merely make such suppositional estimates himself, but had shown that Nozzolini's principle required him, Nozzolini, to maintain that such foolish estimates, because nearer the truth, are better than others that are not foolish, and had predicted that Nozzolini could not hold fast to his principle. When one disputant employs against another the *reductio ad absurdum* of his views, it would be strange if the other could avoid its effect by simply demanding that his opponent must not introduce absurdities. But Nozzolini, rejoicing in this method of rebuttal, went on and likewise threw out all Galileo's use of other quantities, in which there is no gain or loss, on the ground that he had to consider only the problem as originally proposed, which concerned itself only with values, forgetting that he himself had been the first to introduce other quantities with his reference to the butchers.[7]

Beside these rejoinders, Nozzolini makes three or four new arguments, which are left for us to tackle. One is an appeal to a more technical forum than that of the butchers— to the fact that when two or more official appraisers of

[7] But he would allow no reciprocity on Galileo's part. His incomprehensible argument referred to in Note 3 was ignored by Galileo. He now repeats it, says it is the best, and he will stand on it alone, and claims the victory because it was not refuted (p. 75).

some property in litigation differ, the arithmetic mean between their appraisals is everywhere, and always has been, the one legally employed (p. 71B).  To this no different reply need be made than to that other appeal to other people's opinion and practice.    Then he cites the case of archery and of similar contests (rifle practice to-day) to hit a mark, and finds confirmation of his view in the fact that the person who comes nearest to the bull's-eye (whether above or below it) is considered the best shot.    Here he cites a case where practice is theoretically correct ;  but it is one essentially different from the subject of estimation.  For estimates can err only above and below the mark, while shots may go astray in any direction ;  and, which is the principal thing, the limits of error in shooting at a mark are the same in all directions, being nowhere ;  but in estimation, while the errors above may go up without limit, the errors below quickly reach a limit when the estimate falls to zero.    Now, it is precisely the geometric mean which sends its extremes at each remove in such wise that  the lower one never reaches zero before the upper one reaches infinity ;   whereas the arithmetic mean sends its lower extreme down to zero as soon as the upper extreme has reached double the quantity aimed at.    Thus, while in errors of archery and the like, where the limits on both, as on all, sides are equally distant, or distanceless, the arithmetic mean is the proper one to employ ;  in errors of estimation, which have a definite lower limit and none but an infinite upper limit, the proper mean, in theory at least, whatever be the practice, is the geometric mean.    Again, he argued that if a thing weighs 60 pounds, and both the wrong estimates are on the same side, say 55 and 50, we give the decision in favour of the nearer, and so judge them by their distance ;  and just so, when the estimates are on opposite sides, say 55 and 70, we should do the same, judge them by their respective distances,—or rather, he asks, why should we not ?  Here he is betraying signs of dotage ; for he overlooks that Galileo had over and over again shown why we should not ;  and even in the first supposition, where the difference of proportion accompanies the difference of distance (wherefore Nozzolini cannot rightly say we use the latter only), there is a begging of the question in his assertion ;  for, according to the opposite contention, we can no more correctly say that an error of 20 per cent. above the truth is twice as great as an error of 10 per cent. above the truth, than we can rightly say that the latter

error is equal to an error of 10 per cent. below the truth. Lastly, in a supplementary letter, he introduces a principle never heard of before and never again since, which he can hardly word clearly, but which is something to the effect that in all questions of greaterness we go from the greater to the smaller, and in all questions of smallerness we go from the smaller to the greater ; and so we must do in a question of estimates, that is, we must go, as he before contended, from the true value in opposite directions to the greater and to the lesser extremes ; whereby he seeks to convict Galileo of having to maintain that $\frac{1000}{100}$ is $= \frac{10}{100}$, whereas, of course, Galileo had maintained that $\frac{1000}{100}$ is $= \frac{100}{10}$. It is a principle invented as an afterthought for the purpose of supporting the conclusion desired ; and it deserves no further consideration.

Nozzolini had commenced with great moderation, but he ended by betting his eyes that the geometric proportion has nothing to do with measuring errors of estimation. He displays the signs we expect in a man who argues on the wrong side—of floundering into greater and greater absurdities as he persists in his course.

## II.—AVERAGES AND ERRORS.

So lengthy a dispute over so useless a problem may seem puerile. But a problem may be useless in itself, and yet be profitable as an exercise in reasoning. The controversy we have been reviewing was not a mere logomachy : it may be viewed as a methodological essay. It can serve both for a model and for a warning, the former provided by Galileo, the latter by Nozzolini. And there are indications to-day that the one as well as the other is needed. Even some of its direct results have had to be rediscovered in our own time, as will be seen more than once ; and it is a pity they were ever forgotten. Hence that controversy by no means deserves the contempt with which Todhunter alluded to it, as having been a waste of time, and still being without " any scientific interest or value." [8] To draw the lessons that may be learnt from it, let us first inquire into the nature of averages, and then into the nature of estimates and their errors, after which we may extend the investigation to more practical problems.

### The Criteria of Averages.

In drawing a mean or average the most definite purpose we have is to simplify our quantitative conception of things by obtaining one figure that can take the place of two or more and give the same result as they, numerically speaking. This may be illustrated in four examples. (1) Of two provinces with equal territory, if the population of the one

---

[8] *History of the Theory of Probability*, p. 6. He referred to Libri's *Histoire des Mathématiques en Italie*, Vol. IV. pp. 288-9, as saying that Galileo was occupied with this question (" *non encore resolue* ") a long time. Galileo, though he spent some time over methods of estimating distances and elevations with the eye, did not linger long over this problem of measuring estimates themselves. It is, indeed, a pity that he did not reply to Nozzolini's last letter. He might then have pointed out the distinction between shooting at a mark and guessing at a value, and have saved succeeding generations from error through ignoring it.

is 10,000 and of the other 20,000, this is the same for the state as if the population of each were 15,000 ; for the *sum* of the population is 30,000, which is distributed over two provinces.[9]  Thus the mean population of any country is measured by the arithmetic mean or average.    It is not even necessary to know the individual items : if we know the whole population of a country and the number of provinces (or of square miles) in it, we get the average by dividing the former by the latter.    (2) If the population of a country has increased at irregular rates, but with the result that it has increased 100 per cent., *i.e.*, has doubled, in 100 years, the same result would have been obtained if it had increased regularly at the rate $\sqrt[100]{2} = 1.007$, or 0.7 per cent. every year ; for whatever may have been the 100 individual rates of increase (or even of decrease at times), their *product* is two, and this product is likewise obtained by 1.007 multiplied by itself 100 times.    Here, then, the average increase is the geometric average of the individual increases ;  and again it is not necessary to know the individual items, all that is needed being to know their number and their product. (3) If a merchant sells one article for 190 dollars and another for ten, the net result is the same as if he sold each for 100. Here again the mean to use is the arithmetic.    If the articles were each worth (or their proper selling prices were) 100, his gain and loss would be equal and would nullify each other ; that is, the result to him would be the same as if he neither gained nor lost on either article, as indicated by that mean. (4) If a merchant alternately gains and loses 90 per cent. in his successive business transactions, in every case venturing his whole business capital, what mean gain or loss will have the same result ?  Here the model is twofold, according as he begins by gaining or by losing.    It may be thus tabulated :

| 100 | 190 | 19 | 36.1 | 3.61 | 6.859 | 0.686 |
| 100 | 10 | 19 | 1.9 | 3.61 | 0.361 | 0.686 |

It is seen that of every second venture the result is the same, being an ever growing loss, and the intervening divergent figures rapidly converge.    As there are two lines of develop-

[9] The same *numerically*, be it remembered ;  for in other respects it may be much better for a state that its population be distributed in one way than in another.   And so when we compare unequal territories by their average populations (on equal territories), we are doing no more and no less than what we do when we compare equal territories simply by their populations. The same remark applies to the other examples.

ment, it is impossible for the mean to have more than a recurrently equal effect, while in the intervals it must lie between the divergent figures. Evidently the arithmetic mean is inappropriate for this case ; for the arithmetic mean between 190 and 10 is 100, which indicates that the capital would remain 100 after every venture, contrary to what does take place. The geometric mean between 190 and 10 is 43·6, indicating a loss of 56·4 per cent. ; and now if every transaction showed such a loss, the result would be as follows :

| 100 | 43·6 | 19 | 8·284 | 3·61 | 1·57 | 0·686 |

Thus at every second remove the figures are the same as before, and in the intervening stages they are at the geometric mean between the other two ; but ultimately this series approaches the same limits as do the others, and before long it will be practically indistinguishable from them even in the intervals where it does not exactly coincide with them. It may be objected that the supposition should be that the merchant divides his capital and alternatively gains and loses on the halves, oppositely in each case. Then the figures would be :

| 50 | 95 | 9·5 | 18·05 | 1·805 | 3·43 | 0·343 |
| 50 | 5 | 9·5 | 0·95 | 1·805 | 0·18 | 0·343 |
| 100 | 100  19 | 19 | 3·61 | 3·61 | 0·686 |

Again at every second period the result is the same as in the preceding suppositions, but in the intervals the results are the arithmetic means between the first divergences ; which gives a jerky progress, with two stages on a level, followed by steeper falls. If it be further objected that at every venture the merchant should re-divide his capital into two halves, whereupon his capital would remain intact at 100, in accordance with the arithmetic mean ; the reply is that this returns the case to the second example, of the merchant making equal gains and losses on articles of equal value, which of course gives the same result however many times it be repeated. But we have desired to make a distinct example. And so, while this case is kept as first stated, the proper mean to use on it is the geometric.[10]

[10] Also the harmonic mean may be the proper and necessary one in certain cases. Thus in the ancient problem of the fountain (found in the Greek *Anthology*, XIV. 135—the statement here is slightly altered) the time required to fill the basin through two pipes together which singly half-fill it in different times, is obviously a mean between these times ; and calculation shows

In all these cases the right choice of a mean is demonstrative, and not subject to dispute ; for the kind of mean chosen is determined by its yielding a known result. But in many cases there is no such definite criterion at hand. Then the choice may be determined by the resemblance of the subject to some demonstrative case ; which resemblance may be found in the conditions set to, or the observed behaviour of, the extremes, or the terms in general. The mean to be chosen in the case of estimation can be determined only in this way. Thus it is to be noted that Nozzolini always identified this problem with our third example, and Galileo assimilated it to something like our fourth ; for Nozzolini converted it into a question of gain and loss, and Galileo treated it as a question of business capacity. But Galileo went further, and pointed out an underlying distinction between these two kinds of subjects for averaging. In the one there are no limits on either side of the true position, and in the other there is a limit at zero on the one side and on the other side the only limit is infinity. In the former, any figure above or below the true position may have a corresponding figure on the opposite side equidistant from the true position ; and we may add, if there is a limit on the one side, there must be an equidistant limit on the other. In the latter, the limits zero and infinity are equally related to the true position only as extremes in a proportion, $o : m = m : \infty$ ; wherefore similar equality must be sought in all the intermediate terms that correspond to each other, $x : m = m : y$. This we may generalise, and say the former condition is suitable for the arithmetic mean or average, the latter for the geometric.[11]

that it is the harmonic mean. Note that the lower extreme (the shorter time in which one of the pipes half-fills the basin) can never be quite as small as half the mean (the time in which it in conjunction with the other pipe will fill it) ; and so the harmonic mean is applicable to the case, while the unsuitableness of the arithmetic and geometric means is easily seen by using extravagant suppositions, as of one pipe taking a day to do what the other will do in an hour. Here the only purpose the harmonic mean can serve is to facilitate the calculation the next time we are confronted with such a problem.

[11] We might likewise work out some corresponding condition for the harmonic mean. This is omitted as not needed. Some hints of it may be found in *The Measurement of General Exchange-Value*, Chapter VIII. For the criteria see there especially pp. 245–8.

Thus in our fourth example it would have been possible to suppose the merchant at every other venture to gain 110 per cent., or any higher percentage, but we cannot suppose him in the intervening ventures to lose any more than 100 per cent., remembering that we are making these suppositions as indications of his business capacity ; for this cannot sink below nothing, and if he were supposed to lose 100 per cent. of his capital, he would have no capital to start with for another venture, and the series would come to an abrupt end after the first or second venture.      But in the third example there is no such limit.      We might change the figures and suppose that the merchant sells the hundred-dollar article for 300, making a gain of 200, and then we can ask how he should dispose of another hundred-dollar article to make an equal loss, or to make up to the purchaser for his first unjust sale.      Here, to be sure, there is no price that will satisfy the question ; but it is easy for us to answer it by saying he should give away the second article and 100 dollars to boot.[12] There is nothing ridiculous here.      But if any one has made an error of 200 dollars in estimating at 300 an article worth 100, it would be ridiculous to say he could equal this on the other side by estimating the same or an equally valuable article at 100 dollars less than nothing.      There are no more limits to losses than to gains ; for a man can get himself into the position of having less than nothing by running into debt.      But there is a limit to errors of estimation below the true value, as no man can estimate a valuable article at less than nothing.

We may note, further, that Nozzolini always obtained his result by slightly altering the problem, converting it from one of estimation into one of gain and loss.      In many cases slight changes in the way of stating a problem, or in the attitude assumed toward it, or in the use expected to be made of it, will involve a change in the solution.      Another instance may be cited.      Consider two cities each of 10,000 inhabitants, and suppose 9,000 migrate from the one to the other.      The one has lost 9,000, the other has gained 9,000.      The one has been reduced from 10,000 to 1,000, the other has been increased from 10,000 to 19,000.      Which of these changes is the greater ? and what average are we to use in measuring them ?      Now, if these two cities are both in the same country, and we have in mind the effect of the changes, not upon the cities themselves, but upon the country at large, it is evident

[12] If in this transaction the price be *conceived* as $-$ 100, the arithmetic mean of the sales, $\frac{1}{2}$ (300 $+ -$ 100), is still $=$ 100.

that in mere numbers the country at large has neither gained nor lost ; and so, no effect being produced, the changes have nullified each other, and therefore are equal. So conceived, the case comes under the first example above given and resembles also the third, and the arithmetic mean is the right one to employ. But suppose again that the cities are in different countries, or in different ages, and, without migration between them, the one has dwindled from 10,000 to 1,000 and the other has grown from 10,000 to 19,000. We now have to consider, not the combined effect upon any whole of which the cities are parts, but the effects upon the cities themselves, individually and independently. And now we cannot say the changes are equal. Equal changes in opposite directions would be that of a city dwindling from 10,000 to 1,000 and that of a city growing from 1,000 to 10,000 ; but this last is equal to that of a city growing from 10,000 to 100,000 ; and as the second city in the supposition has grown from 10,000 only to 19,000, it has not grown as much as the first has dwindled. Therefore the mean by which to judge these changes is the geometric.

A little over a century after the controversy we have been reviewing, Daniel Bernoulli in his *Specimen Theoriæ Novæ de Mensura Sortis* [13] and Buffon in his *Essai d'Arithmetique morale* [14] made a similar use of the geometric mean. Bernoulli, to whom the distinction between " moral " and " mathematical " expectation is due, and Buffon argued to this effect : if two men of equal fortunes stake half their fortunes on a throw of dice, the one loses *more* than the other gains. Now, from the standpoint of the wealth of the country of which they are citizens, there is evidently no gain or loss, numerically considered ; and so, when we draw the average wealth of the parties after the play, we must use the arithmetic mean. But from the point of view of the individuals the case is different ; for the one who has lost half his fortune, compares his present state with his previous state, and his previous state was twice as good ; and so does the other, but his previous state was two-thirds as good as his present state, and therefore he looks upon (or feels) his gain as a gain of only one-third, since what he has gained is the third of his present state.[15] In other words, the esteem-

---

[13] Arts. 5 and 13. This work was written in 1730–1, and published in the *Commentarii Academiæ Petropolitanæ*, 1738, Vol. V.

[14] Sections 11–13. This work was written in 1760, and first published in 1777.

[15] Todhunter pointed out that Buffon at least did not do full

value of the money which a man owns or commands has an inverse relation to its amount : as his fortune increases, the esteem-value of his money falls, and as his fortune decreases, the esteem-value of his money rises, and it rises faster than it falls at every arithmetically equal remove from the same initial position ; so that losses are estimated in money of greater esteem-value than are gains. If the losses and gains (or the stakes in a play) are small compared with the person's fortune, the difference is imperceptible ; but it becomes apparent when they are comparatively large.[16]   Hereby is

justice to his case. The loser has lost a fortune equal to his present fortune, while the winner has gained only one-third his present fortune. Todhunter says that, representing the original fortune by $a$ and the stake by $b$, Buffon estimated the gain by the expression $\dfrac{b}{a+b}$ and the loss by $\dfrac{b}{a}$, whereas the loss should be estimated by $\dfrac{b}{a-b}$ : *History of the Theory of Probability*, p. 345. Bernoulli and Buffon have been followed, among others, by Laplace, who introduced the distinction between a person's " moral " and " physical " fortune, *Théorie analytique des Probabilités*, 1812 (in his *Œuvres Complétes*, 1847, Vol. VII. pp. 474–88, cf. p. xxii. of the Introduction) ; Lacroix, *Traité élémentaire du Calcul des Probabilités*, 2nd ed., 1822, p. 127 ; Quetelet, *Lettres sur la Théorie des Probabilités*, 1846, let. viii. ; Fechner, who brought the principle under Weber's law, *Psychophysik*, 2nd ed., Vol. I. pp. 236–8, Vol. II. pp. 549–50 ; Jevons, *Theory of Political Economy*, 1871, p. 154. On the other hand Venn in his *Logic of Chance*, 2nd ed., 1876, pp. 135–7, 408–12, does not accept this reasoning. He points out that the pleasure of the game has been overlooked, as also the temperament of some people, who are bored with the monotony of mediocrity and delight in excitement. This explains why gambling always has been, and probably always will be, popular, in spite of the loss of " moral fortune " which attends it ; but it does not disprove that loss. That loss is a payment for the amusement obtained— and it may be an overpayment through ignorance on the part of the players. Cajori, in his *History of Mathematics*, 2nd ed., p. 223, remarks that Bernoulli's theory has become classic, " but no one ever makes use of it." So much the worse for those who might but do not.

[16] In all the above by a man's fortune must be meant, not merely his actual capital, but in addition to that the capitalised value of his earning power. A man possessing property worth 20,000 dollars, who earns 10,000 a year, may lose his whole capital without feeling much loss. As Bernoulli pointed out,

explained the well-known phenomenon that people grieve more over losses than they rejoice over arithmetically equal gains ; and cynics are mistaken when they sneer at this. In the case of our two gamblers, individually considered, as the one has lost more than the other has gained, the net effect is a loss ; and this is shown by drawing the geometric mean of their final positions. Bernoulli calculated the loss of what Laplace, who accepted his reasoning here, called their " moral fortune " at 13 per cent., because $\sqrt{150 \times 50} = 87$ nearly, and this lies 13 below 100.[17] Bernoulli, Buffon and Laplace reasoned as Galileo had done.

Different ways of viewing a problem, however, need not lead to the use of different averages. An instance in point is that of drawing the average height of the men of a country or of a race. Our purpose may be merely to compare the stature of the men of one country or race with the stature of the men of other countries or races, or of men to-day with men of earlier times. Such a use of the average has most resemblance to that in the first model above given ; for we may suppose that all the various heights of all the men of one country are to those of another as they would be if all the men of each country were uniformly tall and yet together had the same combined height that they actually have. We may conceive of all the men in each country standing on top of one another, forming two upright columns, and then, provided their numbers are equal, the comparison of their heights is the comparison desired, and the same result is given by comparing the arithmetically averaged heights of

even a beggar has some fortune, which is the capitalised value of the gifts he annually receives. Without this correction it is wrong to say that 10,000 is as great a sum to a man owning 20,000 as is half a million to a man owning a million. The former is likely to be much better able to replace his loss than the latter. Hence Franklin was not quite right when he said, in 1768, in connection with universal suffrage, that " the all " (evidently meaning all the property) of a poor man is " as dear to him " as the all of a rich man is to him, *Works*, Spark's ed., Vol. II. p. 372 (and again so reported by J. Baynes in Sir Samuel Romilly's *Memoirs*, 3rd ed., Vol. I., p. 447). Franklin's dictum was controverted on this ground by Horne Tooke in his *Letter to Lord Ashburton*, quoted in the *Edinburgh Review*, Vol. XXXI. p. 183.

[17] Going back to the point of view of the country at large, we now see that it too has suffered a loss, as it has got a worse distribution of its wealth. But this cannot be shown in the merely numerical calculation of its wealth.

the individuals in the columns ; and if the numbers are unequal, the proper correction is made by using these averages.    In the case of races, the total heights are beyond our powers of ascertaining, and then we must use a number of each race sufficiently large to serve as samples, and must average them arithmetically.    Most statistical averages are drawn for this purpose of comparison, especially in time series, to see whether there has been improvement or deterioration, with a view to finding causes, whereby we may check the latter and forward the former.    And whenever such combination of the items is possible, the arithmetic average is properly indicated.    But an entirely different purpose for drawing anthropometric averages was invented by Quetelet.    Quetelet wished to obtain the type of the men of a country or of a race.    This type is conceived as something really existing, and if we could know it otherwise (by revelation, say,) we might use it to see which kind of average fits it, and thereby determine our choice of the average proper for this case.    But as we can learn it only by first drawing an average, the kind of average to use is not determinable in this way.    We are thrown back on the other criterion provided us by Galileo and Nozzolini in their controversy. If the heights of men behave like estimates, the geometric average should be used.    If they behave like gains and losses, the average to use is the arithmetic.    The heights of men have one resemblance to estimates : they never fall to zero ; but they are very different from estimates in another respect, in that they have never been found to exceed a certain figure, which is hardly more above the general run of human stature than the lowest recorded limit is below it. They come under the class, therefore, of things with limits equidistant on the opposite sides ; and the proper average to use for finding the typical or normal height of man would still seem to be the arithmetic.

The last example discloses a defect in part of the reasoning above employed, which needs to be corrected.    We have found the resemblance of these later examples to our definite models to consist in the behaviour, so to speak, of the terms. If they behave like the arithmetic terms around their average, or if they behave like the geometric terms around their average, the arithmetic or the geometric average is to be chosen accordingly.    But this implies that we know the average already, which sometimes is contrary to the supposition ; for in other cases, notably those of change or variation, knowledge of the total effect brings them under one of

the models.    Now, in the last example we supplied for the average what we there called " the general run of human stature."    This is something knowable when we have a large assortment of data.    In mathematics it has been called, in our language, the " mode," which is a meaningless word. The German Fechner named it the " *dichtester Wert*," or the " thickest value " ;  and it is a pity we have not some equally significant word—the " mean of greatest thickness " has been suggested, which, however, is too clumsy.    But in the absence of a better word we must put up with this one.    The mode, then, is found by arranging the quantities under treatment according to their sizes, and if, as is generally the case when great numbers are collected, they cluster most thickly around some one position, this is the mode.    Now, it may be found that the numbers in the different groups of quantities thus arranged taper off equally, or about equally, on each side of the mode.    This is the behaviour of arithmetic terms, and indicates that the arithmetic average of them should be employed.    But it may also happen that they taper off more thinly and spread out more above than below the mode ;  and this is the way geometric terms around their average behave.    In theory, for arithmetic terms, if they be grouped in numbers that fall within certain equal intervals— for instance, if the heights of men be measured only by inches, so that all heights falling say between 4 feet $11\frac{1}{2}$ inches and 5 feet $\frac{1}{2}$ inch are grouped together ;  then, on the supposition that the group just mentioned is found to be the mode for some people or race, the intervals equidistant above and below it will contain equal numbers of items.    But this will not be the case with geometric terms.    To form groups containing as many items above as the corresponding groups contain below the mode, the upper intervals must be gradually enlarged as they recede from the mode.    On the lower side, then, the intervals must be contracted, in order that they may still be equal geometrically.    For instance, on the supposition that the stature of men followed the geometric arrangement, the mode remaining the same as before, if the lower interval around 2 feet 6 inches has shrunk to 2 feet $5\frac{3}{4}$ inches—-2 feet $6\frac{1}{4}$ inches, then the corresponding upper interval would have its position and bounds set thus : its lower bound $= \dfrac{25}{2\frac{25}{48}}$ $=$ 9 feet 11 inches, and its upper bound $= \dfrac{25}{2\frac{23}{48}} =$ 10 feet 1 inch approximately ;  so that this interval would be twice as far removed and four times broader.

What we have been doing, is to study the *dispersion* of the data, and we find that there may be an arithmetic and a geometric dispersion.[18]   The stature of men is found to conform rather to the arithmetic dispersion ;  and this is the reason why the arithmetic average is the proper one to use in averaging the heights of men (also for Quetelet's purpose).   Were it possible to measure our mental capacities quantitatively with any exactness, Galton's investigations of human intelligence indicate that it more likely follows the geometric dispersion ;  for his researches disclose that men of the greatest genius are to men of mediocre ability as these are to idiots or witless men.[19]   If this be so, the proper average for averaging degrees of intelligence would be the geometric.   To prove that actuality conforms with theory in this subject, an infinite number of data would have to be collected.   In practice not even a very large number, mathematically speaking, is ever collected.   The data actually collected, therefore, can be expected to conform only approximately with the theoretical requirements.   If the mode in our collections agrees more closely with the arithmetic average, this average should be employed ;  and the geometric, if it agrees more closely with the geometric. It may, however, happen that the mode can be only roughly outlined, and then if (as is likely) the two averages fall close together, the accuracy of the discrimination may be greatly impaired.   It may also happen (which is not unlikely with a small assortment of data) that the dispersion is so irregular that no mode is discernible, or there may be no mode to discern, according to the nature of the subjects.   Then this criterion has failed us, and we must do the best we can without any, employing analogy with known cases where we can, or else in practice adopting the most convenient average, which is the arithmetic.   Instances are known in which two or more modes appear, which may be taken to indicate that the data belong to different species, *e.g.*, to different

[18] By examining the behaviour of harmonic terms we might find also a harmonic dispersion.   But with this we are not concerned.

[19] Cf. *Hereditary Genius*, end of Chapter III.   Galton expresses the matter in an arithmetic progression ;  but his whole work shows that the progression is rather the geometric.   When he wrote that work in 1869, he had not knowledge of geometric dispersion.   Cf. William James :   " As the genius is to the vulgarian, so the vulgar human mind is to the intelligence of a brute," *Psychology*, II. p. 348.

races of men, when their stature or other measurements are studied.[20]

Doubtful cases are of frequent occurrence in statistics. Statisticians frequently have to deal with ratios. Now, it might be thought that because, when we deal with *differences*, we must use the arithmetic average, therefore when we deal with *ratios*, we should always have to use the geometric average. But this is not necessarily so. If ratios express variations—increases or decreases,—then the geometric average must be used. Even in the case of human stature, if our object is to measure its growth, as in the case of a boy who has grown irregularly in five years from 4 feet to 6 feet, the proper average of the individual rates of increase is the geometric. In other cases perhaps there is a presumption in favour of the geometric average ; but necessary evidence is lacking.

To study ratios, we must first of all distinguish them from mere fractions, since they are expressible in the same form as fractions. Fractions proper express the comparative number of parts in a common whole. Ratios express the comparative relations between distinct quantities. With fractions there is no difficulty. If we have occasion to use the arithmetic average between 6, 8, and 9 inches, we still must use the arithmetic average between these quantities when expressed as fractions of a foot, viz. $\frac{1}{2}$, $\frac{2}{3}$, and $\frac{3}{4}$ ; and the result will be the same. But in social science there is need, among other things, of expressing the ratio of the annual deaths in a country's population to that population, thereby obtaining what is called the country's death-rate, which is usually taken to be the numerator when the denominator has been reduced to 1,000. Here the deaths are not parts of the population ; wherefore the death-rate, though expressed in the form of a fraction, or as so many thousandths, is not a mere fraction, but is a ratio proper. Again, in anthropometry there is need of expressing the ratio of the breadth of the head to its length, which ratio, in a similar manner generally expressed as so many hundredths, is called the cephalic index. Such cephalic indices as 70, 80, 90, and the like are not parts of 100, but they are relations to 100. Now, evidently the death-rate must lie

---

[20] So Quetelet, *op. cit.*, p. 143. Adolphe Bertillon has been able thereby to distinguish two races in the Department of Doubs, near the eastern border of France, *La Théorie des Moyennes en Statistique*, in the *Journal de la Société de Statistique de Paris*, 1876, pp. 289–93.

between 0 and 1, or as usually expressed between 0 and 1,000 ; for the deaths cannot sink below zero and cannot rise above the population. As for the cephalic indices, it is conceivable that they might rise above 100, as we know no reason why the breadth (from right to left) of somebody's head should not be greater than its length (from front to back) ; but no such cases are found, nor are cases found of the cephalic index falling much below 60. Thus there are limits above and below, within which the death-rate and the cephalic indices range, the latter having only empirical limits, and the former having additional extreme limits. These cases, therefore, belong to the class where the proper average may be either the arithmetic or the geometric, according to the position of the mode in a large collection of such figures. Statisticians have not yet taken the trouble to find which of the two is the proper one, and perhaps many of their subjects will not yield clear results. They have by common consent adopted the arithmetic average, as being the easiest, like the butchers to whose example Nozzolini appealed. Moreover, at least in the case of cephalic indices, they have been bothered by having two different methods at their disposal.

To calculate the cephalic indices of thousands of skulls, and then to draw the arithmetic average of them, is a laborious task. To shorten the labour, many anthropologists have adopted the method of adding together all the breadths and all the lengths, and of then drawing the ratio between these as the one common or average ratio of the lot. This, however, gives a slightly different result from the preceding, sometimes above, sometimes below, though it may sometimes agree. Bertillon has given instances. Using only two skulls for simplicity, he pointed out that if the one is 180 mm. broad and 200 long, and the other 112 broad and 160 long, the index of the former is 90 and of the latter 70, between which the arithmetic mean is 80 ; but if we divide the sum of the breadths by the sum of the lengths, 292 by 360, and multiply by 100, we obtain for the mean index a higher figure, 81·1. Again, if the one has the measurements 144 and 160, and the other 140 and 200, the indices are likewise 90 and 70, with the mean at 80 ; but if we divide their combined breadths 284 by their combined lengths 360, and multiply by 100, we obtain for the mean index a lower figure, 79. Lastly, if the measurements be 144 and 180, and 136 and 193, yielding the same indices and the same mean, the latter method also yields the mean

index 80. The proper method, Bertillon asserts, is to average the indices separately, in the first way ; for then, he says, their dispersion can be studied.[21]    It would seem more correct to say that we should obtain as many indices separately calculated as we can, and by arranging them in the order of their magnitudes study their dispersion for the purpose, first of all, of discovering which is the proper average to draw. It might be found that the proper average is not the arithmetic, which Bertillon has simply assumed, but the geometric. In the instances cited the geometric mean between 90 and 70 is 79·3, which perhaps is the best of all.[22]    But in practice the geometric average, even with the aid of logarithms, is a very troublesome one.   Therefore even if it were the theoretically correct one, it might be advisable in practice to employ one of the other two methods, since these in practice give results very slightly different not only from each other but from the geometric average. Still, as the true one, the geometric average could serve as a norm for deciding between the other two ; for " mathematicians and anthropologists," says one of the most recent writers on statistics, " differ as to which method is more correct."[23]   What, then, should be done ?

Letting $\frac{a'}{b'}$, $\frac{a''}{b''}$, $\frac{a'''}{b'''}$, . . . represent the cephalic indices

(in their original form) of $n$ number of skulls, we express the three methods in these three formulæ :

$$\frac{1}{n} \left( \frac{a'}{b'} + \frac{a''}{b''} + \frac{a'''}{b'''} + \ldots \text{ to } n \text{ terms} \right), \qquad (a)$$

$$\frac{a' + a'' + a''' + \ldots \text{ to } n \text{ terms}}{b' + b'' + b''' + \ldots \text{ to } n \text{ terms}}, \qquad (b)$$

$$\sqrt[n]{\frac{a'}{b'} \cdot \frac{a''}{b''} \cdot \frac{a'''}{b'''} \ldots \text{ to } n \text{ terms}}, \qquad (c)$$

The first is the simple arithmetic average of the several ratios, and the third is the simple geometric average of them. The second is of a rather recondite nature. It can be shown to be two things and very approximately a third thing.

[21] *La Théorie des Moyennes en Statistique*, pp. 304–5.
[22] Indeed, the cephalic indices 90, 80, 70, etc., have a resemblance to the lesson-marks of school-children, which, we have seen in Note 5, ought to be averaged geometrically.
[23] Zizek, *Statistical Averages*, W. M. Persons's translation, New York, 1913, p. 19.

(1) It is the weighted arithmetic average of the ratios with weights according to the figures in the denominators; for it is equal to

$$\frac{\frac{a'}{b'} \cdot b' + \frac{a''}{b''} \cdot b'' + \frac{a'''}{b'''} \cdot b''' + \ldots}{b' + b'' + b''' + \ldots}, \qquad (d)$$

which is the formula for such a weighted average. (2) It is the weighted harmonic average of the ratios with weights according to the figures in the numerators; for it is equal to

$$\frac{a' + a'' + a''' + \ldots}{\frac{a'}{a'/b'} + \frac{a''}{a''/b''} + \frac{a'''}{a'''/b'''} + \ldots}, \qquad (e)$$

which is the formula for such a weighted harmonic average. And (3) it can be shown, both by theory and by trial, to be in ordinary cases very approximately equal to this formula,

$$\Sigma \sqrt[\sqrt{ab}]{\left(\frac{a'}{b'}\right)^{\sqrt{a'b'}} \cdot \left(\frac{a''}{b''}\right)^{\sqrt{a''b''}} \cdot \left(\frac{a'''}{b'''}\right)^{\sqrt{a'''b'''}} \ldots} \qquad (f)$$

which is the weighted geometric average of the ratios with weights according to the geometric means of both the numerators and the denominators in the individual ratios.[24] It is of course also equal to

$$\frac{\frac{1}{n}\,(a' + a'' + a''' + \ldots \text{ to } n \text{ terms})}{\frac{1}{n}\,(b' + b'' + b''' + \ldots \text{ to } n \text{ terms})},$$

which is the ratio of the arithmetic average of the numerators to the arithmetic average of the denominators: something which need not further concern us, any more than the fact that the third is equal to

$$\frac{\sqrt[n]{a' \cdot a'' \cdot a''' \ldots \text{ to } n \text{ terms}}}{\sqrt[n]{b' \cdot b'' \cdot b''' \ldots \text{ to } n \text{ terms}}},$$

[24] It should be noted that if we change the figures in some of the ratios, or in all of them differently, but without altering the ratios themselves (substituting for instance $\frac{2}{4}$ for $\frac{1}{2}$, or $\frac{xa}{xb}$ for $\frac{a}{b}$, and the like, but variously), this mere change of the figures does not affect the result in the first and third formulæ, but it does affect the result in the second, since it alters the weights of the ratios differently. Again, of course, if all the figures are changed in the same proportion, no effect on the result is produced even in the second; for weights are relative.

which is the ratio of the geometric average of the numerators to the geometric average of the denominators. The first three facts have generally not been recognised by mathematical statisticians, if even by mathematicians. Fechner, for instance, has named the second formula the " summary mean," as if it were something distinct from the other means, although he perceived that in it the ratios with greater figures have greater weight.[25] It is extremely important that the true nature of this expression should be recognised.

Now, in the case of the cephalic indices, when the second method is employed, leaving aside the fact that its advocates are employing, quite unconsciously, the harmonic average of the ratios weighted according to the breadths of the heads, and also (approximately) the geometric average of them with weights combining the breadths and the lengths, but confining our attention to the arithmetic average, we see that they are really using this average of the ratios weighted according to the lengths of the heads! Evidently there is no rhyme or reason for using such weighting as this. But if the arithmetic average itself is a wrong one for the case, perhaps it with a wrong weighting may give a better result than it would give with the right weighting, since two

[25] *Kollektivmasslehre*, 1897, pp. 164, 353, 357. On pp. 359–61 Fechner compared the second formula with the third ; but it did not occur to him to compare it with *f.* Likewise Walras, making no inquiry into weighting, distinguished a formula for index-numbers of prices like the second, as a " multiple-standard combination," from methods of averaging price-variations arithmetically or geometrically, as if it were something distinct, *D'une Méthode de Régularisation de la Valeur de la Monnaie,* pp. 15–16 of the offprint from the *Bulletin,* Vol. XXI., of the Société Vaudoise des Sciences Naturelles, 1885. And many other economists have done the same. Even the mathematician Cauchy, in his *Cours d'Analyse,* 1821, Vol. I., 1st theorem, although he demonstrated that the above second formula is *an* average of the given ratios, yet did not point out what kind of an average it is. This was done in *The Measurement of General Exchange-Value,* pp. 504, 511, 519 (resting the proof of the approximation of formula *f* to formulæ *b, d,* and *e* on the fact that the geometric average is always less than the arithmetic and greater than the harmonic, between the same terms with the same weights). But already Professor Irving Fisher had noticed, in connection with price-variations, the identity of *b* with *d* only, *The Rôle of Capital in Economic Theory, Economic Journal,* December, 1897, pp. 517, 520.

wrongs may neutralise each other, while one wrong alone remains wrong. We shall have occasion for this remark later, whether it be applicable here or not. What statisticians should do, therefore, is to find whether the simple geometric average is the right one for these cases, and if it is, then let mathematicians find which of the two practical methods gives results in the long run closer to the results given by the geometric average ; and if it should turn out to be the second method, the second method should be adopted, in spite of its curious weighting, for practical purposes. Or if it turns out that the simple arithmetic average is the right one, even then the second method may continue to be used, as the most convenient, and not appreciably erroneous for ordinary practice. At all events statisticians will then know exactly what they are doing.

Thus the theoretically true method, even though not adopted in practice, may serve as a guide to practice. It is only by first getting the theoretically true method that we can know how near to it our practical methods come. The theoretically true method, therefore, is important, even though we do not employ it in practice.[26]

[26] In De Morgan's *Essay on Probabilities*, 1838, pp. 156–7, occurs a curious passage with regard to the death-rate and similar ratios (which he speaks of simply as " fractions "). Different calculations of the death-rate may give different results, which need to be averaged. He supposes three such results to be $\frac{12}{20}$, $\frac{13}{22}$, and $\frac{17}{30}$, or in decimals 0·6, 0·591, and 0·567. The simple arithmetic average of these is 0·586, and their average by the second formula above, which he did not recognise as an arithmetic average with weights according to the denominators (the deaths), is 0·583, being smaller merely because it happens that the smaller ratios have received the greater weights. Then he asserts that neither of these is right, and the right (or most probable) result is obtained by multiplying each numerator and denominator by the denominator and dividing the sum of the new numerators by the sum of the new denominators, which in this case gives 0·581, being still smaller for the same reason enhanced. Now this, though he did not recognise it, is the arithmetic average with weights according to the squares of the denominators (the deaths). No reason is apparent for such weighting, any more than for the weighting in the second average. Note that if the calculations of the deaths and of the population were made by six different operators, the ratios can be combined in six different ways (*e.g.*, one other would be $\frac{12}{22}$, $\frac{13}{30}$, and $\frac{17}{20}$), and each of the three

The upshot of this inquiry into averages is the discovery of two criteria for their use, the second of which has some subdivisions.   The first criterion is a known result, with which the mean or average must agree.   The second criterion comes into play when the result is the object sought, and it consists in a resemblance between the ways the terms behave and the ways terms behave around arithmetic and geometric averages when these are known.   Here there are three different ways of behaviour : either (1) the terms extend both above and below without any conceivable or assignable limits, and then the arithmetic average must be used ; or (2) there is a definite lower limit at or above zero and no upper conceivable or assignable limit, and then the geometric is the right one to use ; or (3) the terms either cannot, or at least so far as we find in experience do not, fall below some figure, and do not, or cannot, rise above some other figure, which two figures serve as limits either necessary in the nature of the things or found in practice to exist ; and then, when enough data are provided for examining their orderly dispersion, if their mode is found nearer to the arithmetic average, the arithmetic average is the proper one to use, and if their mode is found nearer to the geometric average, the geometric average is the proper one to use ; or else this criterion fails, and we are left to be guided by analogy or by convenience.

It is important to note that the examination of the dis-

averages just used would most likely give a different result in each case.   But if we use the simple geometric average, it is indifferent which combination be used, the same result being given in all.   This result for the above figures is 0·585.   Another peculiarity of De Morgan's treatment of this matter is that he offers the method he so dogmatically recommends, as " the method of least squares " (and so again in his article on *Least Squares* in the *Penny Encyclopædia* and in his *Theory of Probabilities*, § 124, in the *Encyclopædia Metropolitana*, Vol. II., p. 452), although it is the first average (the simple arithmetic) which always gives the least sum of the squares of the residual errors in such cases, while the method he recommends gives, in this case, the greatest sum of all !   What his method does, is to give the least sum of the squares, not of the errors simply, but of the errors multiplied by the denominators.   This, however, may be argued for on the analogy of the reduction of errors in measuring, say, a mile ; for here we do not draw the arithmetic average of the errors expressed as fractions of a mile, but of the errors expressed absolutely (or as those fractions multiplied by their denominators).   But it is questionable whether this analogy holds.

persion for this purpose is needed only in these last ambiguous cases. If the data behave in either of the first two ways, the average is determined by the second criterion with as much certainty as it is by the first criterion, and there is no need of reference to the mode in their dispersion. Take for instance the second model above given : if all the 100 rates of increase that have contributed to the doubling of the population in a century were known, we could arrange them in the order of their magnitudes and find the number of rates falling in each magnitude, and it might turn out that they exhibited the geometric dispersion ; but if they did not, we should still use the geometric average for finding the average annual rate of increase. And it is quite possible that with so small a total number of terms as 100, their dispersion would not contain a clearly defined mode, or one nearer to the geometric than to the arithmetic average ; but we have no concern whatever it may happen to be. The same certainty as to the appropriateness of the geometric average appertains to the fourth model, and likewise to the case of estimates—the estimates we have so far been dealing with,—although this belongs under the second criterion. It is possible that if a thousand persons should independently estimate the length of a given line, we might find their estimates to fall in geometric dispersion. But it is not necessary to wait for this experiment before deciding that the geometric average is the proper one for averaging estimates of this sort. The appeal to the mode in the dispersion of the data is only a last resort in dubious cases—and of course only in case of sufficient importance to make it worth while.

## ESTIMATES AND OBSERVATIONS.

Coming back to estimates and their errors, we must note first of all that the word " estimate " is used in two different meanings. In the one meaning it refers to valuations mentally made, or drawn at sight or by the aid of other senses— in short, to guesses, at least to such as are made with care. In the other meaning an estimate is a measurement, or the result of many measurements, or in astronomy they are observations, since astronomical observations are frequently accompanied by measurements. And the errors committed in these two kinds of estimates may be distinguished as errors of estimation and errors of observation, the former being errors of guesses, the latter errors of measurements.

Errors of estimation are the ones we have been dealing with, and Galileo has shown that they must themselves be measured or estimated geometrically. But errors of observation are entirely different : they must be measured arithmetically. For, although the former can and do go much further above than below the true position, the latter need not deviate more above than below. Surveyors in measuring and re-measuring a given distance, say about a mile, let them measure it no matter how carefully, never get results that are not slightly discrepant, within limits of at least a few inches ; but there is no reason to suppose their errors are more above than below the true distance ; [27] wherefore when they average their measurements in order to eliminate their errors as much as possible, they very properly use the arithmetic average. These considerations show that while estimates proper are to be averaged geometrically, measurements are to be averaged arithmetically. If you *measure* a line with instruments many times, you should use the arithmetic average of the measurements ; if you *estimate* a line by the eye many times, you should use the geometric average of the estimates.

Whether Galileo had occasion to draw an average between different observations of the same quantity, and which average he used, it is not easy to gather from his works ; but there is one passage which indicates that he would not have hesitated to use the arithmetic average. In his *Dialogues*, published in 1632, near the beginning of the third, he maintains that in adjusting various discordant calculations of the distance from the earth of a new star,

[27] In measuring a line with a ruler our errors, so far as caused by the expansion or contraction of the ruler under changes of temperature, may be greater in excess than in defect, because expansion and contraction are subject to the geometric scheme, since there is an absolute zero temperature and no superior limit. But our position in the heat scale is so high compared with the variations to which our rulers are exposed, that the difference in their expansion and contraction above and below the standard chosen is infinitesimally small and quite negligible. For a similar reason, if whenever we look at a thermometer we first guess the temperature, our guesses would probably be no greater (appre-ciably) above than below the true figure. Here we should be practically aiming, so to speak, at a single mark, the other end of the scale being out of our experience ; wherefore the arithmetic average should be used of such guesses. Thus guesses also are of a two-fold nature.

based on twelve observations in different latitudes, the most probable result is obtained by retaining all the observations and by employing the smallest possible emendations that will bring them into agreement ; and he based the last on the consideration that the errors of the observers were as likely to be above as below the truth, and as great on the one side as on the other, and more likely small than great.[28] All these assumptions underlie the principle of using the arithmetic average in correcting observations. But whatever his practice or want of it, his successors soon had to employ an average in reducing their observations, and they rightly adopted the arithmetic average. And they used the arithmetic average so much, and their example was followed by civil engineers and others to such an extent, that it came to be forgotten that there were or might be cases that call for the geometric average.[29] The suitableness of the geometric average for estimates proper seems to have been re-discovered in 1879 by Galton. Galton started from Weber's law, which had twenty years before been thoroughly elaborated by Fechner. Because nerves react, from a certain initial position (after crossing the threshold of consciousness) in arithmetic progression as the stimuli increase in geometric progression, Galton argued that our estimates of weights by handling bars, or of shades of colour by looking at them, should be averaged geometrically. He remarked that " the ordinary law of frequency of error based on the arithmetic mean," which " asserts that deviations in excess must be balanced by deviations of equal magnitude in deficiency," requires that " if the former be greater than twice the mean itself, the latter must be less than zero, that is, must be negative " ; but " this is an impossibility in many cases, to which the law is nevertheless applied by statisticians with no small success, so long as they are content to confine its application within a narrow range of deviation." He would, therefore, extend the use of the geometric average to sociological subjects as well,[30] thus reversing the course

[28] *Opere*, Vol. II. pp. 302 A and B, 311, and 312 A.

[29] The want of knowledge of the distinction between errors of observation and errors of estimation is especially conspicuous in the opening passage of Chapter VII. of De Morgan's *Essay on Probabilities*.

[30] *The Geometric Mean in Vital and Social Statistics*, in *Proceedings of the Royal Society*, 1879, XXIX. pp. 366–7. Galton won precedence by a narrow margin ; for, apparently in independence of him, the next year in Italy, in a paper on *Il Calcolo*

of history, since Weber's law of vital phenomena had received its first hint from Bernoulli in the field of sociology and economics. Galton was followed by Professor Edgeworth, who a few years later wrote : " The estimates which different persons (or the same person at different times) might make of a certain weight would be likely to err more in excess than in defect of the true weight, and in such wise as to render the geometric mean of such a series of estimates the proper method of reduction."[31]   But the re-discovery has not yet made its way into general recognition and acceptance.[32]

*dei Valori Medii e le sue Applicazioni statistiche*, in the *Archivio di Statistica*, Anno V., 1880, also as a pamphlet, Rome, 1883, pp. 31–5, Messedaglia reviewed the Galileo-Nozzolini controversy, and accepted Nozzolini's solution for the purpose of measuring gain and loss in buying and selling and Galileo's for the purpose of measuring the precision of estimates, since estimation can err with equal facility more in excess than in defect ; and immediately thereupon he noticed the distinction between this proper geometric way of measuring estimates and the usual arithmetic way of measuring observations, and explained it by the subjectivity of the former and the objectivity of the latter, and also by invoking Fechner's exposition of Weber's law.   But there Messedaglia left the subject.

[31] *Memorandum* to the first Report of the Committee on Index-Numbers of the British Association for the Advancement of Science, published in the Report of the Fifty-seventh (1887) Meeting of the Association, 1888, p. 283.   Again, in his *Memorandum* to the second Report of the Committee, in the Association's Report of the Fifty-ninth (1889) Meeting, 1890, p. 201, alluding to a hypothetical error in certain data of 25 per cent., he adds in a parenthesis " (an error which, if possible *in excess*, is almost inconceivable *in defect*)."   He could not, however, find proof of geometric dispersion in examination-marks, although he expected it there, and suggested the use of the geometric average for averaging the marks of different subjects, *The Statistics of Examinations*, in the *Journal of the Royal Statistical Society*, September, 1888, pp. 607, 625.

[32] An example illustrating how the distinction continues to be overlooked is furnished by Merriman, who in his text-book on *The Method of Least Squares*, § 138 (still in the latest revised edition, 1911), sets a problem about 1,600 persons *guessing* the contents of a vessel, and after giving the (presumably arithmetic) average of their guesses expects his students to solve a question in connection therewith in the same way as all the other problems in the book, all the others being about observations or measurements.

We, however, having seen abundant reason provided by Galileo and Nozzolini, for this distinction, are in a position to accept it unreservedly. And now some further study of the difference between the behaviour of observations, dispersing arithmetically, and of estimates proper, dispersing geometrically, needs to be made.

Galton gave this subject to the mathematician Dr. McAlister to work out mathematically, which was done, and the result was published in sequence to Galton's own paper, under the title of *The Law of the Geometric Mean*. Whether Dr. McAlister performed the task completely and satisfactorily, does not appear to have been examined by other mathematicians. It cannot be examined here. All that can be done is to note the difference between the curve of observations and the curve of estimates.
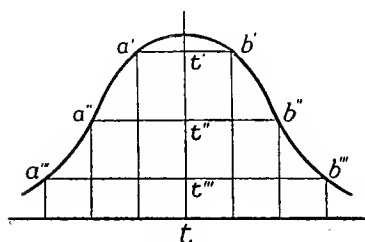


FIG. 1.

For by representing observations and estimates as abscissæ and their frequency as ordinates, a curve may be drawn; and such a curve, in the case of observations, is well known because of its bell-shape or resemblance to a *gendarme's* hat. It is more or less elongated according to the accuracy of the observations. A common form of it is shown in Fig. 1. It is a symmetrical curve around the axis drawn vertically upon the point $t$, which represents the true magnitude, so drawn that the horizontal lines relate to each other thus :

$$at = tb ;$$

for these lines represent, on each side of the axis, the amounts of the errors, which in the case of errors of observations are assumed to be equal on both sides of the true magnitude. This magnitude $t$ is a distance from some point on the base line to the left, which need not be indicated in the figure, because it has no influence on the proportions of the curve. But in drawing the curve of estimates it is necessary to introduce this point (with its vertical line), because all the estimates have reference to it. The curve of estimates will therefore be like that drawn in Fig. 2, on the supposition that the errors in defect of the true magnitude are somewhat

like those in the preceding case. The difference now shows itself most strikingly in the curve on the right side of the axis, since on this side are represented the errors in excess, which extend as much beyond the true magnitude as the errors in defect fall short of it, proportionally to $o$, having the relation

$$oa : ot = ot : ob.$$

It will be noticed that, in Fig. 2, if the magnitude which is estimated be enlarged by pushing the zero point further to the left, then, on the supposition that the curve on the left remains the same by the estimates becoming more accurate, the curve on the right will be drawn in toward
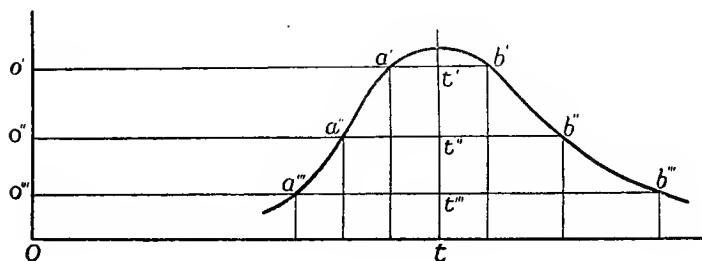


FIG. 2.

the axis, and if that magnitude be increased infinitely by removing the zero point altogether, the curve on the right will become in all its points equidistant from the axis as the curve on the left, reducing the whole to the same symmetrical shape as in Fig. 1; for now we have virtually changed the estimates into observations around a single point, having done away with any reference to the line $ot$, and having only the position of $t$ to consider.[33] In observations, it is true, we always measure some magnitude $ot$, but on examining our instruments in the neighbourhood of $t$, as also in comparing our observations afterward, we have no use for the line $ot$; but in forming and comparing estimates we must always keep the reference to the line $ot$. An estimate, in fact, is always a ratio, since it has no meaning except in relation to the true magnitude (determined by measurement); but a measurement has independent validity, since

[33] Cf. above, note 27.

it can be corrected only by another measurement.[34]   Common parlance bears out this distinction ; for we mostly dissociate errors of measurement from the magnitude measured, but in errors of estimation we always relate the errors to the magnitude estimated.[35]

We have seen above that *in theory* the dispersion around the mode, which in the figures is obviously the axis, gives equal numbers of items at corresponding intervals on each side, the difference between the arithmetic and geometric dispersions consisting in the fact that in the former the intervals are equal in breadth, but in the latter the intervals above (to the right) are broader than the intervals below (to the left) and progressively so as they depart from the mode.[36]   In the case of observations we must suppose

[34] Although a measurement does not give the exactly true magnitude, it is so much more precise than an estimate, that its indication is always conceived as the true magnitude in comparison with the estimate.   Similarly, however, measurements made with a rough instrument can have their average controlled by measurements made with a more precise instrument.   " We can appeal from kitchen scales to an atomic balance," as Professor Edgeworth has well said, in the *Journal of the Royal Statistical Society*, September, 1888, p. 601.

[35] We may add the curve of harmonic dispersion :
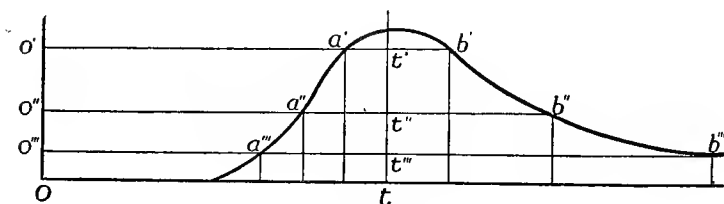


FIG. 3.

This must be drawn so that
$$oa : ob = ot - oa : ob - ot ;$$
whence, if *oa* is given, $ob = \dfrac{oa \times ot}{2oa - ot}$.   The figure gives the proportions correctly, but cannot vouch for either side separately.

[36] In the calculus of the probabilities of errors of observation use is made for various purposes of the areas contained between the base and vertical lines and the curve, because the observations are equally distributed on both sides.   As this is not the

(1) that errors are equally likely to be above as below the true magnitude, so that in the long run there will be an equal number of them on opposite sides of the true position, and (2) that equal errors are equally likely on the opposite sides. We must believe these two things because we have no reason for the contrary, and because the supposition is supported by experience of the fact that in large numbers of observations so-called residual (or reputed) errors do so behave around the average of the observations.[37] In the case of estimation the first supposition still holds : we have no reason to believe that estimates will be more numerous on one side of the true position than on the other. But in estimates the second supposition must be interpreted as referring to geometric equality, or else be stated in the form that greater errors in excess are equally as numerous as the (arithmetically) lesser errors in defect to which they (geometrically) correspond. As for the first supposition, which is unequivocally common to them both, this involves that the mode must coincide with what is called the " median," or midmost item in the series, and both the mode and the median must coincide with either the arithmetic average, in arithmetic dispersion (of observations), or with the geometric average, in geometric dispersion (of estimates). This, remember, is the theoretical requirement. But the theory itself requires that this condition shall be fulfilled only in an infinite series. Short of infinity, the theory declares that it will be very closely obeyed in a very large series or collection of data, and that the discrepancy of fact from the theoretical condition will increase as the number of the data collected decreases. In practice, therefore, we cannot expect the arithmetic average in its proper subjects or the geometric average in its proper subjects in every particular case exactly, and in some cases even approximately, to conform to the mode and still less to the median ; or, which is the same statement reversed, we cannot expect the mode or the median to coincide exactly with the proper average of the subjects, or even with each other. Now, when they do not coincide, it is these averages, the arithmetic

case in geometric dispersion, in which the estimates are progressively thinner from the left to the right, no such use can be made of the areas in calculating the probabilities of errors of estimation.

[37] Residual errors are not true errors ; but their behaviour in conformity with the theory of true errors adds inferential strength to the theory.

and the geometric, each in its proper subjects, which are the summarising quantities, the single ones that may stand for and take the place of the manifold quantities, because yielding the same numerical results as they yield, or the indicative quantities, as yielding most nearly what we wish our calculations to yield (this last may be affirmed on the analogy of the first) ; and it is not the mode or the median that can serve these purposes of the averages. The mode, and the median too, can serve, in dubious cases (and only in such is their service needed), to help us find which of the two averages is the correct one to use ; and then their work is done. They cannot be properly used to oust and to replace those other averages. Yet in practice, since the mode and the median diverge as a rule but slightly from the proper averages, they can, and have been, used for convenience in their place (just as the arithmetic average has been used in place of the geometric) ; but with them, on the whole, the saving of labour is so small, and they, especially the median, are so capricious, that it is poor policy on the part of statisticians to recommend their use.[38]

[38] That is, as substitutes of the other averages ; for they may serve other purposes. It may be questioned whether even mathematicians have done well in choosing the median of the errors to measure the " probable error " in observations, instead of the arithmetic average of the errors. The fact that in a series of observations there are as many errors beyond as within a certain error does not import that in a repetition of observations the errors will fall equally around that one rather than around the arithmetic average of the previous errors ; and the fact that mathematicians have chosen the arithmetic average of all the observations to get the one most probably true, would seem to require them to use the same average to get the error most likely to be exceeded and to be fallen short of. Otherwise they might as well have chosen the median of the observations as giving the most probable true value ; for they could have argued thus : The observations hitherto made are equally numerous on each side of a certain magnitude, therefore it is probable they will continue to be equally numerous around that magnitude ; but it is probable that observations fall in equal numbers on opposite sides of the true magnitude ; therefore that (median) magnitude is most probably the true magnitude. But they do not argue this way in this case. Then why should they argue so in the other ? But of course, as they have almost unanimously chosen to use the median for the " probable error," too little advantage could be obtained to make the change advisable.

Still another point with respect to observations and estimates, especially the former, needs to be cleared up before we pass on. When dealing with the criteria of averages, we could not speak of the criterion of the average proper for observations without anticipating the distinction here drawn between them and estimates ; and now having drawn this distinction, there is need of reverting to the criteria of the averages proper in them. The criterion of the geometric average proper for estimates—that it must be used of them because they have a lower limit and no upper limit and consequently disperse more above than below their mode—has never been disputed, since it has rarely been considered at all. But in the case of observations, the criterion enjoining the arithmetic average because their limits are equidistant from their mode and consequently they disperse equally above and below, has apparently been thought too simple ; for something else has sometimes been put in its place.

Errors of observation have been found to disperse themselves around their mode or median like the chances of black and white balls coming out together in various combinations, from half-and-half up to all-of-one-kind, from an urn containing an infinite number of black and white balls in equal proportion—a condition which can be produced by putting the ball back in the urn after every draw, as then there is no limit to the drawing. Mathematicians have investigated the law of such chances, and have found them to agree with the numbers in one of the lines of Pascal's arithmetical triangle and with the co-efficients in the development of a binomial raised to a power equal to the number of balls in the combination ; and Gauss discovered the formula of the curve following the dispersion of these combinations around the central most probable half-and-half combination. And the law of such combinations, transposed, has been called the " law of error." It is a pity it has been called so, because it really is a law of chances, and it is a law of error only in a different way. It is a law *of* chances, subjectively, belonging to chances, as resulting from the consideration of chances ; while it is a law of error only objectively, as governing errors, because these, when accidental and unbiassed, can be conceived as resembling even chances. It is, too, a law of only one kind of errors, whereas we have seen that there are two kinds, and the other is not subject to this law. The law of the one kind of errors was originally worked out for astronomical observations, which in practice

conform to the law of chances ; for as measurements they in the long run fall equally and equidistantly on both sides of their mode, and their small residual errors are more frequent than their large residual errors.    The latter fact is not essential to the propriety of applying to them the arithmetic average ; but the former fact is.    The fact that the limits of errors of observations are equidistant on opposite sides of the true position, is the whole and sole justification for the application of the arithmetic average to observations and their errors.

Now, this conception of the subject has not always been entertained ; for another justification or so-called proof of the propriety of the arithmetic average for this purpose has been thought to be supplied by what is known as a " doctrine, principle, or method of least squares." Observations of a single magnitude are reduced by the simple arithmetic average.    But two or more magnitudes often have a definite relation to each other, as for instance the internal angles of a triangle, which together must equal two right angles, so that when surveyors measure the three angles of such a figure they know what the sum of their measurements ought to be.    Or, again, two ways of measuring the same thing ought to agree, as when the elevation of a place above the sea is measured from another place the elevation of which has been measured, and then is measured from a second place the elevation of which has been measured, these two sets of measurements ought to agree.    Such measurements (individually made with several observations reduced by their arithmetic average) hardly ever do agree or fit together exactly.    Therefore they need to be adjusted by correcting, obviously not one of them alone, but all of them.    In such adjustments an average cannot be used. But the arithmetic average has three properties—this may be said notwithstanding that one of them belongs to it only as it agrees with the median.    (1) The sum of the differences between the arithmetic average and all the items above it is equal to the sum of the differences between it and all the items below it (or the algebraic sum of all the differences is zero).    (2) The sum of the differences between the median (and consequently the arithmetic average so far as it agrees with the median) and all the items, taken absolutely, is a minimum.    (3) The sum of the squares of the differences between the arithmetic average and all the items is a minimum.    Each of these properties of the arithmetic average (or the median) can be used in adjustments ;

for the aim may be to make the differences required by the corrections to behave in any one of these three ways. Reversely, when any of these methods of adjusting the measurements of several magnitudes is applied to the many observations of a single magnitude, it reduces to the arithmetic average (or in one case to the median) ; so that the arithmetic average is safe in any case (except that in the second it depends upon its agreement with the median). Legendre and Gauss discovered the third method. Before them the first had mostly been used, even by Laplace, who, after Legendre's publication of the third method in 1806, adopted and elaborated it. Which is the best, or the most probable ? A good illustrative example has been given by Whewell. He supposed a quantity known to increase uniformly (as the rising of a star) to be observed at equal intervals, and the measurements to be, say, 4 1/8, 12, 14 1/8. These must be reduced to an arithmetic progression, and he instances three, viz., 6, 10, 14, and 4, 9, 14, and 5, 10, 15. In the first the differences between the measurements and the corrected figures (that is, the reputed or residual errors) are such that their algebraic sum is zero ; in the second, their absolute sum is the least of the three ; in the third, the sum of their squares is the least.[39] The first correction implies that the observer has made two large errors and one very small one ; the second, that he has made one very large error and two very small ones ; the third, that he has made one large error and two small ones. Obviously, on the assumption (always required) that all the observations were made with equal care,[40] the last distribution of errors is the most probable one. Here is a conjunction of the most probable result with the least squares. If mathematicians can prove that this conjunction is not accidental here, but is necessary, occurring in all possible cases, they prove the method of least squares. In effect, the method of least squares introduces a more even low range of corrections, rather than many very small ones and a few great ones ; and so it conforms with the assumption that equally careful observations are more likely to contain small than large errors. But mostly the mathe-

[39] *Philosophy of the Inductive Sciences*, 2nd ed., Vol. II. pp. 408–9. The figures are slightly altered.

[40] If the observations were made with unequal care, they would have to be given unequal weights ; and then the three methods require that the sums of the differences so weighted should behave as stated ; which is only an extension of the simple case.

matical proofs have rested upon a mathematical need of converting negative errors or negative observations into positive quantities, which can be done first by squaring them ; which need goes back to the formula of the curve around the Y-axis, since in it there must be a square the root of which can be both positive and negative.[41]  All this is not very convincing.  But all the proofs have presupposed the very kind of dispersion which calls for the arithmetic average.  Herschel criticised Gauss for basing his proof of the method of least squares upon the use of the arithmetic average in reducing observations of a single magnitude, which, he says, is a " a thing to be demonstrated, not assumed."  He then proceeded to give a proof of his own, and based it upon three assumptions, the third of which was the very assumption upon which the use of the arithmetic average in reducing observations of a single magnitude directly rests.[42]  Others have argued that because the method of least squares is applicable to cases to which the arithmetic average is not applicable, and because wherever the arithmetic average is applied, there the method of least squares is *ipso facto* also applied, therefore the method of least squares is the more comprehensive, and the proof of it is the desired and long-missing proof of the arithmetic average.[43]  They forget that either of the other methods likewise comprehends the use of the arithmetic average (at least the first directly, and the second through the median).  The arithmetic average, on the other hand, comprehends all these three methods (or at least two of them).  Each of them, therefore, is only denotatively more comprehensive than it, while it is connotatively more comprehensive than either of them.

[41] The figures given above are curves of *observations* and *estimates*, all positive, and the Y-axis, or zero origin of the abscissæ, is on the left, and the secondary axis erected on *t* is the axis of the true position of the magnitude observed or estimated.  But as curves of *errors* (in which form Fig. 1 is generally given), the axis *t* is the Y-axis, or zero point of no error, the errors to the right being positive and those to the left negative.  It is to fit this conception of the curve that Gauss's formula for errors of observation was invented.

[42] *Essays*, 1857, pp. 397, 398.

[43] " There is no doubt whatever," says Jevons, " that the method of means [*i.e.*, the use of the arithmetic average] is only an application of the method of least squares," *Principles of Science*, 3rd ed., p. 386.  Cf. Venn, *Logic of Chance*, 2nd ed., p. 336 n.

They forget, too, that the rendering of the least squares is only a property of the arithmetic average, so that when they give precedence to it, they commit a *hysteron proteron*, very much as if metaphysicians should speak of a property possessing a substance, or, to use a more homely similitude, as when humorists speak of the tail wagging the dog. The use of the arithmetic average, when it is properly used, according to the criteria examined in the preceding section, is not based upon, proved by, or even confirmed by the method of least squares : [44] quite the contrary, upon it is based, by it is confirmed, and, if it be proved, by it (with something else) is proved the method of least squares.[45]

[44] Merriman, in his *Method of Least Squares*, § 80, says this method " confirms " the use of the arithmetic average, and refers to § 44 as doing so. But that article merely argued back from the method to the average, going over in reverse order the argument in § 41 which proved " the principle of least squares " by resting it on " a fundamental law " proved in § 26 by a demonstration which rests on the primary supposition of equal dispersion, stated in § 24, which is independently the foundation upon which the use of the arithmetic average rests. It is plain that the method of least squares confirms the use of the arithmetic average only as the conclusion of a long chain of arguments confirms an earlier conclusion much more evidently proved already.

[45] Better than that in the preceding note is the statement of H. Jacoby, in *An Elementary Lecture on the Method of Least Squares, The School of Mines Quarterly*, Columbia University, Vol. XXV. (1904), p. 294, that because Legendre's method of adjusting observations of several magnitudes, when applied to the observations of a single magnitude, reduces to the arithmetic average, " this is a further strong addition to the evidence of plausibility attaching to Legendre's theorem." But even so, of course, the plausibility of Legendre's method is not increased over that of the other two (or at least the first), but needs something else to prove it. It looks as if the method of least squares was invented before it was known that the least sum of the squares of residuals is a property of the arithmetic average ; and when this property of the arithmetic average was discovered through the method of least squares, it gave the appearance that this method was primary and provided the basis supporting the arithmetic average in its use in reducing errors of observation. To substantiate this suggestion would require more research than can be given to it by the writer. The subject of averages has received little attention on the part of mathematicians, and still less on the part of the historians of mathematics.

And when the arithmetic average is directly applied, that is, to single magnitudes of any sort—observations or any-thing else,—the fact that it gives the least sum of the squares of the differences between itself and the items has nothing to do with the propriety of using the arithmetic average when it is proper to use it. The principle of least squares is observed when we average the population of a country per square mile. As well say it underlies, supports, and strengthens the use of the arithmetic average in that case, as to say it does so in the case of averaging observations.[46]

It has been necessary to go into this matter so fully in order to clear the ground of misrepresentations. The use of the arithmetic average in reducing observations needs nothing so complex and mysterious as the method of least squares to rest on. And the use of the geometric average for reducing estimates and other things like estimates, also needs nothing

[46] Fechner noticed that the median has the property that the absolute sum of the differences (or the *first powers* of the differences) between it and the items is the least possible ; and the arithmetic average has the property that the sum of the *squares* of the differences between it and the items is the least ; and some other average (the formula of which he could not find, but a method for obtaining it he worked out) has the property that the *cubes* of the differences between it and the items is the least ; and there may be still another whose *biquadratics* give the least sum, etc., etc. Hereupon he inferred that as the arithmetic average is better than the median, therefore the average giving the least cubes is better than the arithmetic, and the average giving the biquadratics still better, and so on. And to get an application for this sequence, he remarked that the median is good enough when the observations are inexact, and the arith-metic average is required for more exact measurements ; where-fore he concluded that in general the averages giving the higher-powered least sums are the proper ones for observations of greater and greater accuracy: *Ueber den Ausgangswerth der kleinsten Abweichungssumme*, in the *Abhandlungen der koeniglich sächsischen Gesellschaft der Wissenschaften*, Vol. XVIII., *Math.-Phys. Classe*, Vol. XI., Leipsig, 1878, pp. 51–3. There is no sense in all this, except that in adjustments possibly the methods using the least sums of those higher powers might smooth off the corrections still better than the method of least squares, although this is not evident, as they might smooth them off too much. Neither Fechner nor any one else has ever shown what the property of yielding the least sum of differences raised to any power has to do with the probability, in reducing observations of the same magnitude, of approximating to the true magnitude.

of that sort to rest on.    Their uses are perfectly simple and plain.

### THE MEASUREMENT OF ERRORS.

The averaging of observations and estimates is one thing, and the averaging of their errors is another ; for of course errors are different from the observations and estimates which contain them.    Yet, as observations and estimates are different from each other, and are to be averaged differently, so the errors contained in them are different, and are not only to be averaged differently, but are to be measured or reckoned differently.    An account, then, of the methods of reckoning errors is necessary.

Errors can be conceived in at least two ways—a geometric and an arithmetic.    The geometric is the one exhibited by Galileo, being involved in his usage.    According to it the definition of error is this : *A (geometric) error in an estimate is the ratio of the difference between the true value and the estimate, the true value being the subtrahend,* (1) *to the true value when the estimate is less than the true value, and* (2) *to the estimate when the estimate is greater than the true value,* or always to the greater figure, whether it be the estimate or the true value.    If we represent the error by $R$, the estimate by $S$, and the true value by $V$, the definition is expressed thus :

$$R = \frac{S - V}{V} \text{ when } V > S, \text{ and } R = \frac{S - V}{S} \text{ when } S > V.\text{[47]}$$

For example, something having the true value of 14, an estimate of 13 contains an error $= \frac{13 - 14}{14} = -\frac{1}{14}$ or $-7\cdot14$ per cent. ; an estimate of 15, an error $= \frac{15 - 14}{15}$ $= +\frac{1}{15}$ or $+6\frac{2}{3}$ per cent. ; an estimate of 7, an error $= \frac{7 - 14}{14}$ $= -\frac{7}{14}$ or $-50$ per cent. ; an estimate of 28, an error $= \frac{28 - 14}{28} = +\frac{14}{28}$ or $+50$ per cent. ; an estimate of 1, an

[47] If only the absolute figures are desired, the first should read, $R = \frac{V - S}{V}$ when $V > S$.

error $= \dfrac{1 - 14}{14} = -\dfrac{13}{14}$ or $- 92 \cdot 8$ per cent. ; and an estimate

of 196 ($= 14^2$), an error $= \dfrac{196 - 14}{196} = +\dfrac{182}{196}$ or $+ 92 \cdot 8$ per

cent.[48]   Here the same or opposite signs indicate the same or opposite sides ; and equality or inequality of percentage, in absolute figures, indicates equality or inequality of error, but *not* proportional inequality (as will be shown later), according to this geometric method of reckoning error, although it may do so according to other methods of reckoning it.

The arithmetic method of reckoning error was similarly outlined for us by Nozzolini, who gave it in two different forms, the one of which is wrong except when it can be reconciled with the other.    The full arithmetic definition of error is this : *An (arithmetic) error is the ratio of the difference between the true value and the estimate, the true value being the subtrahend, to the true value.*   It is represented thus :

$$R = \frac{S - V}{V}.$$

Nozzolini, we have seen, sometimes treated it as simply a comparison of differences, represented thus :

$$R = S - V ;$$

which means that *an arithmetic error is the difference between the true value and the estimate.*   This is correct when we are dealing with errors around only one true value, as then this one true value is the common denominator in all the fractions, and may be dropped.   It is sufficient in practice when the *difference* between the measurement and the true value is the

[48] That this method of reckoning geometric error is correct, is shown by the fact that an estimate of 13 has to the true value 14 the ratio $\dfrac{13}{14} = 0 \cdot 9286$, which being subtracted from unity indicates a diminution of $7 \cdot 14$ per cent., which is represented as an error by $- 7 \cdot 14$ per cent., as above ; and the true value 14 has to an estimate of 15 the ratio $\dfrac{14}{15} = 0 \cdot 9333$, indicating (a diminution of the true value or) an augmentation of the estimate of $6\frac{2}{3}$ per cent., that is, an error of $+ 6\frac{2}{3}$ per cent.   In general, with the symbols above used, for a lesser estimate we have $\dfrac{S}{V} - 1$

$= \dfrac{s - v}{V}$, and for a greater estimate $1 - \dfrac{v}{S} = \dfrac{s - v}{S}$.

main thing (as when engineers dig a tunnel from opposite sides of a mountain and wish the two sections to meet in the middle ; for here it is the absolute difference which concerns them, and not its relation to the length of the tunnel).    But the second definition always rests on the first, which contains it.    With the same example as before of a thing having a true value of 14, the estimates below the true value have the same errors as before, but above the true value the estimates

have different errors, one of 15 having an error $= \dfrac{15 - 14}{14}$

$= + \dfrac{1}{14}$ or $+ 7\cdot14$ per cent., one of 28 an error $= \dfrac{28 - 14}{14}$

$= + 1$ or $+ 100$ per cent., and one of 196 an error $= \dfrac{196 - 14}{14}$

$= + 13$ or $+ 1300$ per cent.    Here, too, the same or opposite signs indicate the same or opposite sides ; and equality or inequality of percentage, in absolute figures, indicates equality or proportional inequality of error, according to this arithmetic method of reckoning error, proper of course only when the arithmetic average is the proper one to use.[49]

Another definition of arithmetic error, however, has been given by Dr. Bowley in his *Elements of Statistics*.[50]    He speaks simply of error, having in mind only the arithmetic method of conceiving of error.    His definition amounts to this : Error is the ratio of the difference between the estimate and the true value, the estimate being the subtrahend, to the estimate.    His formula, with the symbols here used, is

$$R = \frac{V - S}{S}.$$

He adduced the example we have been using.    The true

value being 14, an estimate of 13 contains an error $= \dfrac{14 - 13}{13}$

$= + \dfrac{1}{13}$ or $7\cdot69$ per cent., and an estimate of 15 an error

$= \dfrac{14 - 15}{15} = - \dfrac{1}{15}$ or $- 6\cdot66$ per cent.    He unfortunately

[49] We might also make a definition of harmonic error.    Cf. above, notes 11 and 18.

[50] Third ed., p. 201.    The definition was given in an article of his in the *Journal of the Royal Statistical Society*, December, 1897, p. 856.

neglected to take the precaution of trying his method on extravagant instances. An estimate of 7 would have an error $= \dfrac{14 - 7}{7} = + 1$ or $+$ 100 per cent., and an estimate of 28 would have an error $= \dfrac{14 - 28}{28} = - \dfrac{1}{2}$ or $-$ 50 per cent. Obviously these results are wrong. When an estimate decreases from the true value to half the true value, it has decreased to 50 per cent. of the true value, and its error has grown to 50 per cent. ; and when an estimate increases from the true value to double the true value, it has increased to 100 per cent. above the true value, and its error has grown to 100 per cent. In the former case, an error of 100 per cent. below the true value is evidently the error of an estimate at zero, but, according to Dr. Bowley, this would have to have an infinite percentage ; while an estimate of one, instead of an error of $-$ 92·8 per cent., would have an error $= \dfrac{14 - 1}{1} = + 13$ or $+$ 1300 per cent. This scheme also does not permit of dropping the denominator when dealing with several estimates about the same object. Moreover, the signs are contrary to the usual practice ; for we conventionally treat excess as positive and deficiency as negative.

Dr. Bowley goes on and gives some examples of the uses that can be made of his formula of error. It may be profitable for us to follow him at least through the first. He supposes two trade unions to return respectively 555 and 45 members as out of work when the true figures are 565 and 50, so that the errors are in the first $\dfrac{2}{111}$ and in the second $\dfrac{1}{91}$ according to his definition of error. He desires " the error in the sum " of these estimates ; and he lays down a rule that " the error in an estimated sum [ $=$ a sum of estimates] is equal to the sum of the errors in the parts when each is multiplied by the ratio of the corresponding part to the sum " (p. 203). Accordingly " the error in the sum " here is obtained so,

$$\frac{2}{111} \times \frac{555}{600} + \frac{1}{9} \times \frac{45}{600} = \frac{1}{40}, \text{ or } + 2 \cdot 5 \text{ per cent.}$$

This is not satisfactory. The total estimates are 600 and the total true values 615, so that, according to Dr. Bowley's

way of reckoning error, the total error is $\dfrac{615 - 600}{600} = \dfrac{1}{40}$,

and the two methods tally ; for what is desired is, obviously, the error in the two estimates added together compared with the two true values added together, which is the same as the error in the arithmetic mean of the two estimates compared with the arithmetic mean of the two true values.   But instead of operating directly in this latter way, Dr. Bowley has sought the true result by using only the errors and the estimates. In reality he has been doing something which he does not

state.   The resultant $\dfrac{1}{40}$ lies between $\dfrac{2}{111}$ and $\dfrac{1}{9}$, and so is

a mean of some sort between them ; and what Dr. Bowley is really after is to find the way of obtaining this mean between the two errors, or what kind of a mean it is.   Here we are dealing with oversights, which can hardly err above the truth, but below the truth may be admitted to err in arithmetic progression, so that no question need be raised about the use of the arithmetic mean.   But the simple

arithmetic mean between $\dfrac{2}{111}$ and $\dfrac{1}{9}$ is not $\dfrac{1}{40}$.   Evidently

uneven weighting must be used. Now, the operation performed by Dr. Bowley is the same as this,

$$\frac{1}{600}\left(\frac{2}{111} \times 555 + \frac{1}{9} \times 45\right) = \frac{1}{40} ;$$

and this is nothing but the arithmetic average between the errors with weights according to the estimates.   It is as if

Dr. Bowley argued : there are 555 errors of $\dfrac{2}{111}$ because the

estimate with this error is 555, and there are 45 errors of $\dfrac{1}{9}$

because the estimate with this error is 45, wherefore the weight of the former is 555 and of the latter 45.   Dr. Bowley is one of those who think that uneven weighting is of little importance in drawing averages except when the figures are few and very divergent (pp. 13–18, 205)—and he is more or less right with regard to practice.   The exception occurs here, and uneven weighting is obviously a necessity.   Only Dr. Bowley has kept concealed (or is it that he did not know ?) that he was using a weighted arithmetic average.   But though these two operations agree, it does not follow that

they must be correct. Were only one mistake involved, there would be little likelihood of agreement. There is need of two mistakes to make agreement in error possible. It may be shown that, apart from the wrong use of the signs, which does not affect the case, another mistake has been committed in addition to a wrong measure of error, namely a wrong system of weighting.

The total error, according to the proper arithmetic method, is $\dfrac{600 - 615}{615} = -\dfrac{1}{41}$, or $-2\cdot44$ per cent. ; and the single errors are respectively $\dfrac{555 - 565}{565} = -\dfrac{2}{113}$ and $\dfrac{45 - 50}{50}$ $= -\dfrac{1}{10}$. The problem, then, is to find the weighted arithmetic average between $-\dfrac{2}{113}$ and $-\dfrac{1}{10}$ that is $-\dfrac{1}{41}$. This is given in the following operation,

$$\frac{1}{615}\left(-\frac{2}{113} \times 565 + -\frac{1}{10} \times 50\right) = -\frac{1}{14}.$$

Here the weights are according to the true values (*not* according to the estimates), and this would seem to be the right system of weighting, or at least a better one ; for the errors are not on the estimates, but on the true values. Otherwise the weights would vary with the estimates and might be smaller on the larger object, whereas it seems reasonable to hold rather that the weights should be fixed according to the true values, an error on a larger figure being more important than an error on a smaller, no matter what the estimates may be.[50a] If, however, we think the estimates ought to count in the weighting as well as the true values, there is no tallying between these two methods of reckoning the same total error.

The correct arithmetic process may be generalised. Let the true values be represented by $v_1$, $v_2$, $v_3$, . . . to $n$ terms, and the estimates by $s_1$, $s_2$, $s_3$, . . . to $n$ terms.[51] Then the

[50a] Dr. Bowley himself remarks : " The greater error in the returns of the smaller union has little effect on the total " ; which seems to imply that the effect on the result should be according to the true values. He has said nothing about the sizes of the unions, but only about the sizes of the returns. He must mean here "the greater error in the smaller returns."

[51] There may be a larger number of estimates than of true values, if one or more values have two or more estimates. The true values must then be repeated for every additional estimate.

arithmetically combined arithmetic errors of the totals is

$$\frac{(s_1 + s_2 + s_3 + \ldots \text{to } n \text{ terms}) - (v_1 + v_2 + v_3 + \ldots \text{to } n \text{ terms})}{v_1 + v_2 + v_3 + \ldots \text{to } n \text{ terms}}.$$

The individual arithmetic errors are $\dfrac{s_1 - v_1}{v_1}$, $\dfrac{s_2 - v_2}{v_2}$,

$\dfrac{s_3 - v_3}{v_3}$, and so on. And the arithmetic average of these, weighted according to the true values, is

$$\frac{\left(\dfrac{s_1 - v_1}{v_1}\right) v_1 + \left(\dfrac{s_2 - v_2}{v_2}\right) v_2 + \left(\dfrac{s_3 - v_3}{v_3}\right) v_3 + \ldots \text{to } n \text{ terms}}{v_1 + v_2 + v_3 + \ldots \text{to } n \text{ terms}}.$$

It is evident that this formula is equal to the preceding formula.

When, now, we try Dr. Bowley's scheme on geometric errors, we no longer have such plain sailing. It is impossible to construct geometric formulæ of averages corresponding to the two just given. We must try the two methods on numerical examples. We shall now find agreement only in a special case.

Let us suppose two towers, 500 and 600 feet high, to be estimated as 600 and 500 respectively. The individual errors, geometrically reckoned, are $\dfrac{600 - 500}{600} = +\dfrac{1}{6}$ or

$+ 16\frac{2}{3}$ per cent. and $\dfrac{500 - 600}{600} = -\dfrac{1}{6}$ or $- 16\frac{2}{3}$ per cent.

We have seen that (1) the arithmetic mean between these, which is 0, indicating 0 per cent. of error, or no error, is correct. In effect (2) the geometric mean between the estimates, $\sqrt{\dfrac{500}{600} \times \dfrac{600}{500}} = \dfrac{1}{1}$, indicates that the mean estimate has the ratio to the truth of 1 to 1, showing that it is correct and there is no error; for the equal errors (geometrically reckoned) on opposite sides (geometrically) nullify each other.[52] The same result is obtained (3) if we seek the geometric error in the geometric mean of the two estimates compared with the geometric mean of the two true values,

for this is $\dfrac{547 \cdot 72 - 547 \cdot 72}{547 \cdot 72} = 0$. Evidently the agreement

[52] Here, too, it happens that the comparison of the combined heights and the combined estimates, the two sums being equal, indicates absence of error.

of these three ways of obtaining the desired result will always take place when there is such alternation of two values and two estimates. For, using the same symbols, we have the condition $s_1 = v_2$ and $s_2 = v_1$, and if $v_1$ be the greater value, $s_2$ must be the greater estimate, and the two errors are $\dfrac{s_1 - v_1}{v_1}$ and $\dfrac{s_2 - v_1}{s_2}$, which on substitution of equals become $\dfrac{s_1 - v_1}{v_1}$ and $\dfrac{v_1 - s_1}{v_1}$, between which two (1) the arithmetic mean is 0, indicating no error ; and (2) the geometric mean of the estimates, $\sqrt{\dfrac{s_1 s_2}{v_1 v_2}} = \sqrt{\dfrac{s_1 v_1}{v_1 s_1}} = 1$, likewise indicates no error ; and of course (3) $\sqrt{s_1 s_2} - \sqrt{v_1 v_2}$, whatever be its denominator, is $= 0$ under the condition, for this makes $s_1 s_2 = v_1 v_2$. And if $v_2$ be the greater value and $s_1$ the greater estimate, the result is the same after some mere inversions of the terms. It can be shown also that the condition may be enlarged to the extent of requiring only $s_1 s_2$ to be $= v_1 v_2$, and then all the three methods will agree in indicating the absence of error in the whole result. Under this broader condition the common indication of no error is obvious in the case of the two last methods, and needs to be proved only of the first. We desire, then, to prove that, $s_1 s_2$ being $= v_1 v_2$, the arithmetic mean of the (geometric) errors is zero. Again there are two cases, for (unless all the terms are equal, in which case there is no difficulty) either $v_1$ is greater than $s_1$, and then $s_2$ must be greater than $v_2$ and the geometric means are as above, or $v_2$ is greater than $s_2$, and then $s_1$ is greater than $v_1$, and the geometric errors are $\dfrac{s_1 - v_1}{s_1}$ and $\dfrac{s_2 - v_2}{v_2}$. In the first case the arithmetic mean between the errors is $= \dfrac{(s_1 - v_1)s_2 + (s_2 - v_2)v_1}{2\,v_1 s_2}$, and in the second case it is $\dfrac{(s_1 - v_1)v_2 + (s_2 - v_2)s_1}{2 v_2\, s_1}$. But from the condition we get $s_2 = \dfrac{v_1 v_2}{s_1}$ and $v_2 = \dfrac{s_1 s_2}{v_1}$. By substitution of these equivalents, and with further help from the condition, the formulæ reduce to $\dfrac{s_1 - v_1 + v_1 - s_1}{2 v_1}$

and $\frac{s_1 - v_1 + v_1 - s_1}{2s_1}$, both of which are $= 0$. Further, with three or more values and estimates, subject to this condition (viz. $s_1 s_2 s_3 \ldots = v_1 v_2 v_3 \ldots$), although the second and third methods will evidently continue to agree, the first will not agree with them, and so will be incorrect.

In other cases (the more usual complex ones) the first method fails entirely, and the third agrees only partially with the second, which remains the correct method. Suppose the same two towers, 500 and 600 feet high, to be estimated as 550 and 450 respectively. The first geometric error is $\frac{550 - 500}{550} = + \frac{1}{11}$ or $+$ 9.09 per cent., and the second is $\frac{450 - 600}{600} = - \frac{1}{4}$ or $-$ 25 per cent. The arithmetic mean between these is $-$ 7.95 ; but this does not correctly indicate the geometrically reckoned whole error.[53] The correct result is obtained in the second way, by drawing the geometric mean between the estimates stated in terms of their true values, thus $\sqrt{\frac{550}{500} \times \frac{450}{600}} = \sqrt{0.825} = 0.9083$, or $-$ 9.17 per cent. This can again be obtained as follows. Reduce both the true heights to unity ; the corresponding estimates then are 1.10 and 0.75, and the geometric mean between these is $\sqrt{1.10 \times 0.75} = \sqrt{0.825}$, as before. The reason why the arithmetic mean between the geometric percentages of errors does not in such cases give the right result is because, as already stated, the geometric method of reckoning error does not correctly give the proportion of one error to another. This is easily seen ; for the estimate 121 is known to be twice as erroneous as the estimate 110 of a true value 100 (for 100 : 110 = 110 : 121) ; but the percentages of error, viz. 10 and 21, do not correctly indicate the proportional magnitudes of the errors. What is equal here is the excess of 110 above 100 and the excess of 121 above 110, for $\frac{110 - 100}{110} = \frac{121 - 110}{121}$. Errors are to be compared by analysing them in this way.[53a]

[53] Nor is it correctly indicated by the added heights compared with the added estimates, since $\frac{1000}{1100} = 0.90909$, or 9.09 per cent., the same (so it happens) as the first error alone.

[53a] Cf. *The Measurement of General Exchange-Value*, pp. 250–3.

Hereby another proof of the above result is obtained.    The
estimate of the second true value below 600 that equals in
error the given estimate of the first true value above 500, is
the estimate 545·45, for 550 : 500 = 600 : 545·45.   But the
given estimate of the second true value is still lower, so that

its unbalanced error is $\dfrac{450 - 545\cdot45}{545\cdot45} = -0\cdot175$, or an error

of − 17·5 per cent.    But this error is distributed over two
estimates, the other of which is correct, since it has been
balanced.    This error is the error of an estimate 0·825, when
the true value is unity.    The other estimate, being correct,
is unity.    The geometric mean between these is $\sqrt{1 \times 0\cdot825}$
= 0·9183, as before.    Furthermore, the correct result in the
instance before us, and in all instances in which the product
of the true values is greater than the product of the estimates,
is given by the third method also, which consists in seeking
the geometric error in the geometric mean or average be-
tween the estimates compared with the geometric mean or

average between the true values; for this is $\dfrac{497\cdot48 - 547\cdot72}{547\cdot72}$

= − 0·0917, or − 9·17 per cent., as before.   We may generalise
this, continuing to use the same symbols.   The formula of
the second method, giving the correct result in geometric

percentage of error, is $100 \left( \sqrt{\dfrac{s_1 s_2}{v_1 v_2}} - 1 \right)$; and the last method,

when the product of the true values is the greater, has the

formula $100 \left( \dfrac{\sqrt{s_1 s_2} - \sqrt{v_1 v_2}}{\sqrt{v_1 v_2}} \right)$, which is evidently equal to

that [54]; but when the product of the estimates is the greater,

the formula is $100 \left( \dfrac{\sqrt{s_1 s_2} - \sqrt{v_1 v_2}}{\sqrt{s_1 s_2}} \right)$, so that in these cases

this method does not agree with that.   The two agree, of
course, also when $s_1 s_2 = v_1 v_2$; for this is the condition in the
preceding paragraph over again.   What is here proved of
two values and two estimates can evidently be proved of any
number of values and the same number of estimates.   But
beyond these methods and the cases to which they are
applicable the correct result cannot be obtained by any

---

[54]. For in these cases, with the product of the true values the
greater, the geometric method of reckoning error is the same as
the arithmetic.

assignable mean or average drawn merely between the (geometric) errors, with any weighting suggested by and confined to the figures given. No geometric mean can be drawn between a positive and a negative quantity. If we replace the negative error by its estimate, the geometric mean between 1·0909 and 0·75 ($= \sqrt{1·09 \times 0·75} = \sqrt{0·81}$) $= 0·9045$, indicating an error of $- 9·55$ per cent., which is wrong, though only to a trifling extent.

Thus only in the simple but rare cases where there are two estimates of two values such that the product of the estimates is equal to the product of the values, can we correctly get the geometrically combined result by drawing an arithmetic mean between the geometric errors—for a geometric mean between them is impossible. Here is a difference between the behaviour of the geometric average and of the geometric mean. The geometric mean when it is applicable (between two sets of two figures whose products are equal) gives a correct indication of the result. But the geometric average (between two unequal sets of two figures, or between three or more sets, whether equal or unequal, provided they do not form two equal groups) does not give a correct indication of the result. This breakdown of the geometric average (between many quantities), in spite of the correctness of the geometric mean (between two evenly weighted quantities), will trouble us again.

But there is a problem in connection with errors of estimation involving unequal quantities and uneven weighting, in which the weighted geometric average works as smoothly as does the simple geometric mean. This is a problem in which the different quantities or magnitudes are joined to each other endwise on opposite sides of a single true position, and so form one geometric progression. Such a problem is presented when the estimates of a single magnitude are given and their comparative errors are supposed to be known, and the true magnitude is sought. This is a widening of one of the problems included in the seventeenth-century controversy. The original question did not leave room for variety of weighting. For, as posited, the problem was to apportion the errors of known estimates around a single known value ; and no question of any but even weighting could come in. And when it was altered into the form of deriving the true value from the estimates and the errors, the errors were assumed to be equal, so that it still could employ only even weighting. We may now, however, assume the errors to be unequal, wherefore we shall have to weight the estimates

unevenly. In this connection there is no failure of the weighted geometric average to give the correct result.

To begin with the simplest case, let us suppose B.'s estimate of a certain value to be twice as erroneous as A.'s estimate of the same. This means that B.'s error is twice as great as A.'s, or that it takes two of A.'s errors to equal B.'s one. Therefore if we suppose another person to repeat A.'s estimate (and consequently A.'s error), we have three estimates, two of which are alike ; and the errors of these two together (not necessarily their sum, yet the last being conceived as beginning where the first left off) are equal to B.'s one error. Thus A's estimate must be weighted as two to B.'s estimate as one, that is, *estimates must be weighted inversely according to their errors* (*or directly according to their accuracy*). This rule, of course, holds for this kind of problem, and not necessarily for others. It is different from the system of weighting we found necessary for Dr. Bowley's problem also when applied to estimates. There we found that the weights should be, not according to the estimates or their errors (although those were supposed to be known), but according to the true values ; for that problem involved more than one true value. So again in the problem of reducing observations (and the same in the case of reducing estimates) in order to get the most probable value (the most probably nearest approach to the true value), the different observations are weighted in some proportion to the care with which they were made, and not according to their errors ; and so, if they are all made with equal care, they are weighted evenly, notwithstanding that some may be much more erroneous than others.[54a]

The weighting being as described, if we suppose the two

---

[54a] In most practical cases, if, after the probably truest value has been ascertained in this way, the observations (or estimates) were to be weighted inversely according to their probable errors and the operation of averaging them thus weighted were to be repeated, so little would be gained in accuracy as to make this additional operation not worth the very much enhanced trouble. Theory shows that in an infinite series of observations (or estimates) the weighted reductions would not differ at all from the unweighted ; and when there are many observations (or estimates), the difference between them would be very slight. When the observations (or estimates) are few, it may be questioned whether this second operation ought not to be performed. And yet in these cases, as no precise result can be obtained, it would be difficult to prove that the second operation yields a nearer approach to the true value than the first alone.

errors of estimation to lie on the same side of the true value, there is little difficulty. It is merely a question of geometric progression, and not of geometric averaging, and so does not interest us. We are interested in cases where the estimates lie on opposite sides of the true value, so that the true value, lying between them, is some average between them—between the estimates, not (of course) between the errors. Let us suppose A.'s overestimate of an article to be 200 and B.'s underestimate of it to be 50 ; and the supposition still being that the error in B.'s estimate is twice as great as the error in A.'s, the problem is to find the true value of the article. If we were dealing with observations and the arithmetic average were employed, the answer would be $\frac{1}{3}$ (200 + 200 + 50) = 150, indicating that 150 is the true value. According to this, A.'s error would be + 50 and B.'s − 100, which at once yields the doubleness of B.'s error over A.'s, in absolute differences. Or if we reckon them in percentages, we should have A.'s error $= \frac{200 - 150}{150} = + \frac{1}{3}$ or + 33$\frac{1}{3}$ per cent., and B.'s error $= \frac{50 - 150}{150} = - \frac{2}{3}$ or − 66$\frac{2}{3}$ per cent., likewise indicating the doubleness of B.'s error. But we are dealing with estimates proper, and we have learnt that a lower estimate contains an error greater than its absolute figure or percentage arithmetically reckoned indicates, so that if the true value were 150, B.'s error would be more than double A.'s. The geometric average is $\sqrt{200 \times 200 \times 50} = 126$ nearly, whereby A.'s error is shown to be + 74 and B.'s − 76, or nearly equal in absolute figures ; but as the error of the lower estimate counts for more than the other, it may be twice as great. When we reckon the errors in percentages by the geometric method, we get A.'s error $= \frac{200 - 126}{200}$ $= + \frac{37}{100}$ or + 37 per cent., and B.'s error $= \frac{50 - 126}{126}$ $= - \frac{60}{100}$ closely, or − 60 per cent., which is not double the other, though nearer to being so. But again we have seen that, while equality in this reckoning indicates equality of

error, inequality does not indicate proportional inequality of error. We must now find how to calculate the doubleness and other manifoldness, geometrically reckoned, of one estimate over another on opposite sides of the true value, when their comparative erroneousness is known.

For this we must begin with the easier task of comparing estimates on the same side of the true value. Let $v$ as before be the true value, $s$ a given estimate, and $x$ (which we desire to find) an estimate twice as erroneous, that is, containing an error twice as great. These three lie in geometric progression in the order named, $v$, $s$, $x$, in which $x = \frac{s^2}{v}$. Now, two equally erroneous estimates (with the same symbols accented) on opposite sides of the true value form the geometric progression $s'$, $v$, $x'$, in which $x' = \frac{v^2}{s'}$. In this last any value of $s'$ may be substituted. Hence an estimate $\frac{s^2}{v}$ on one side of the true value is equalled in error by an estimate on the other side $\frac{v^2}{s^2/v} = \frac{v^3}{s^2}$. Therefore, as this estimate $\frac{s^2}{v}$ is twice as erroneous as the estimate $s$ on the same side, and as it is equalled on the other side by the estimate $\frac{v^3}{s^2}$, an estimate $\frac{v^3}{s^2}$ on either side of the true value is twice as erroneous as the estimate $s$ on the other side. Applying to our example, we replace $v$ by 126 and $s$ by 200, and we find in effect that an estimate on the other side twice as erroneous as an estimate 200 is $\frac{126^3}{200^2} = 50$, which agrees with the supposition, and shows that the error − 76 contained in the estimate 50 is twice as great as the error + 74 contained in the estimate 200.

To generalise the problem, we may proceed with the supposition that B.'s estimate is thrice as erroneous as A.'s. Then on the same side the estimates form the geometric progression $v$, $s$, $x$, $y$. Here we already have the value of $x$ in $s$, and by substituting that value for $x$ in the progression $s$, $x$, $y$, and resolving, we obtain $y = \frac{s^3}{v^2}$. If B.'s estimate is

four times as erroneous, we may abbreviate the progression to this, $v$, $x$, $z$, and by substituting the value of $x$ in $s$, we obtain $z = \dfrac{s^4}{v^3}$. We now have sufficient basis for the mathematical induction that the powers of $s$ and $v$ increase by unity with every additional erroneousness supposed. In general, then, on the same side of the true value an estimate $n$ times as erroneous as a given estimate $s$ is $\dfrac{s^n}{v^{n-1}}$. And now, by the same reasoning as before, this estimate $\dfrac{s^n}{v^{n-1}}$ on either side of the true value is equalled in error by an estimate $\dfrac{v^{n+1}}{s^n}$ on the other side; and therefore, to a given estimate $s$ on either side of the true value, the estimate on the other side that is $n$ times as erroneous is $\dfrac{v^{n+1}}{s^n}$.[55] Here, of course, $n$ may stand for a broken number, or fraction, as well as for an integer.

Now, when an estimate $s_2$ is $n$ times more erroneous than estimate $s_1$, these estimates are the same as if we had $n$ estimates like $s_1$, and one estimate $s_2$, or $n + 1$ all told; and the geometric average between these, which yields the true value of the article, is $\sqrt[n+1]{s_1{}^n s_2}$. In effect, the estimate $n$ times more erroneous than the estimate $s_1$ being $\dfrac{v^{n+1}}{s^n}$, the geometric average between these, $\sqrt[n+1]{s_1{}^n \cdot \dfrac{v^{n+1}}{s_1{}^n}}$, is $= v$.

But the extension of this problem to the averaging of three or more estimates again leads to confusion. When the numbers are odd, it is impossible to have equal numbers of estimates on the opposite sides (unless one is correct, when it may be ignored). Even when the numbers are

[55] *E.g.*, a lower estimate three times more erroneous than the estimate 200, when the true value is 126, is $31\frac{1}{2}$; and one four times more erroneous is $19\frac{4}{5}$, while one equally erroneous is $79\frac{1}{2}$. All these form the geometric progression $19\frac{4}{5} : 31\frac{1}{2} : 50 : 79\frac{1}{2} : 126 : 200$. The figures going from 126 to the left represent estimates respectively one, two, three, and four times as erroneous as the estimate 200.

even and the estimates fall equally on the opposite sides, it is impossible to get the proportional weights for more than two of them at a time. But the difficulty is the same in the case of observations, even with the use of the arithmetic average.[56] The fact is, however, that any two alone of the estimates (or observations) are sufficient, with the known comparison of their errors, to indicate the true value, and the indication yielded by any other two, provided their comparative errors have been correctly stated, will be the same. As there is no need of using more than two opposite estimates (or observations), there is no need of incurring the difficulty.

Indeed, this whole problem, like the one cited from Dr. Bowley's book, seems futile, since we never can know the erroneousness of estimates (or of observations) without first knowing the true value itself. The method of solving it can be of use only in cases where we have some reason for believing the errors of two estimates (or observations) to have a certain proportion to each other ; for then the result obtained for the true value would have the same amount of plausibility in its favour. In practice, however, all that we can know is that two or more sets of estimates (or observations) of the same magnitude have different " probable errors," that is, errors of certain amounts equally likely to be exceeded as not whenever other estimates (or observations) are made in the same ways. These probable errors cannot be treated like true errors, as they have a law of their own. Now, to repeat, in averaging many estimates, just as in averaging many observations, of the same magnitude, we must attach to them equal weights, provided they have been made with equal care and with equally good instruments, and independently of one another. But if we possess two calculations of estimates of the same magnitude, the one set made by experts and the other by less experienced persons, we shall probably find less dispersion in the former than in the latter, which would be a sign of the greater accuracy of the former. In the analogous case of observations, when we possess two calculations on the same magnitude, the one set made with instruments of greater precision than the other, the dispersion in each set is used to calculate

---

[56] But if the errors are all equal (arithmetically or geometrically) in pairs, every one on one side having another corresponding to it on the other, there is no difficulty about averaging them all together simply, either arithmetically or geometrically.

the probable error of its result ; and now in drawing the average between the results of two such sets the usual practice is to *weight* them *inversely according to the squares of their probable errors*.[57]   This operation should rest on the assumption that the two sets err on opposite sides of the true position.   Such a distribution cannot be known, but the assumption can be made on the general principle that errors are as likely on the one side as on the other ; and, too, if they do happen to be on the same side, we do not know on which side it is, and so even in this case the chance is as good of decreasing as of increasing the error, while in the case of their happening to be on the opposite sides (equally likely), the chance is improved of decreasing the error.   When there are three or more sets with different weights, they must still be averaged, on the general principle of probabilities, in spite of the fact that this cannot be done in the case of true errors ; but there is no difficulty about averaging any number of equally good sets.   As for estimates, they ought to be treated in the same manner, with substitution of geometric methods.   Only in these cases there is little need of such treatment, since for greater accuracy we had better betake ourselves to measurements.

[57] Cf. Merriman, *The Method of Least Squares*, §§ 63, 88.

## III.—DEVIATIONS AND VARIATIONS.

IF the treatment of errors of estimation were the only use we had for the geometric method, its difficulties would be of little moment. But this is not the case. Characteristic of errors is it that they are deviations from a central position, either known or conceived as the true position. Allied to them are variations, which differ from them only in that variations are changes from one position to another equally good. Deviations are conceived as going in opposite directions from a centre. Variations may meet and cross one another while moving in opposite directions between distant positions. But just as deviations may have their starting points separated, so variations may have their starting points brought together, and then they move like deviations in appearance, but with the important difference just noted, that in deviations the central position is viewed as their proper resting place, while in variations we know no reason why the position from which they start is better than the position in which they end.

Deviations and variations are common subjects in statistics. Here the central position is viewed, for deviations as a type or norm from which the objects depart, and for variations merely as a common starting point artificially made. And the purpose of drawing averages may be, for deviations to find the type or norm, and for variations, not to find the common starting point, which is known, but to find the average amount of the deviations, in order to learn how much the objects have varied all taken together or as a whole. And all the deviations and variations dealt with in statistics fall into at least the two kinds we have been discoursing upon. There are deviations and variations which depart from a starting position (roughly indicated for the former by the mode or median, and afterward more exactly by the average adopted) about equally on both sides; and there are others which depart from it more above than below it, as reckoned in the usual arithmetic way. Deviations and variations that disperse equally on the opposite sides are mostly deviations of single quantities; those which disperse more above than below are mostly

ratios between independent variables, although this is not a necessary or proved distinction. It is obvious that the arithmetic average is the right one to use of deviations and variations that move in the former way, and the geometric of those which move in the latter way.

A great many of the deviations and variations dealt with in statistics, like the heights of men already noticed, behave in the former way, and are fit subjects for the arithmetic treatment. In ignorance of any other kind of errors, statisticians have generally likened all the deviations and variations they deal with to errors of observation, the principles of which have been elaborated to their hand, as we have seen, by astronomers and mathematicians. Astronomers have rightly ignored any other kind of error, since their science is confined to observations, in which the errors move in arithmetic dispersion. Statisticians likewise have little use for errors of estimation, but they are concerned with other objects of which the deviations or variations may move in geometric dispersion. Hence, in confining themselves to the use of the arithmetic average, they have done what astronomers have not done : they have applied it to cases to which it does not belong. On account of their usual practice hitherto of using only the arithmetic average, it is difficult as yet to enumerate many of the subjects which ought to be treated geometrically. But at least one subject immediately suggests itself.

## AXIOMETRY.[58]

The subject which immediately suggests itself is very prominent in statistics and economics. It is the deviations of prices at a later period from what they were at an earlier period, when they had a level which it is desirable to maintain—and it is desirable to maintain any price-level that is more or less settled. It is, however, the whole price-level that it is desirable to maintain, not the prices of individual commodities, which should be allowed to vary according to the rule of demand and supply, the former determined by people's needs, the latter by their ability to produce the commodities. Therefore price-changes are variations rather than deviations, and their individual positions in the price-level which it is desirable to maintain, are merely the starting

[58] This term is proposed for the measurement of general exchange-value, especially of money.

points of their possible and even desirable variations. We often speak of averaging prices, as if prices were single objects (like the heights of men) ; which is perhaps one reason for the tendency to average them arithmetically. But prices are not single objects ; they are ratios—the ratios of the values of commodities to the value of the money-unit. Strictly speaking, there is not even a level of prices, as an absolute object, like the level of the ocean in a calm. A level of prices exists only in comparison with another level of prices at another period, which other level of prices may be the same with, or different from, the level at the first period.[59]   Jevons has said that " there is no average of prices at any one period " ; [60] this is too broad a statement. But the truth is, we do not really average prices in axiometry ; and the only meaning of an average of prices would be to determine the amount of preciousness of commodities, or the relation of their value to their weight or bulk—an object of little interest.[61]   In axiometry what we really do is to average the variations of prices. The common starting point, if we have constructed it, from which we begin to measure the *deviations* of prices, is purely arbitrary.   At any period of time all commodities have various prices ; but we can find of the commodities various amounts that have the same price.   Then at the next period these amounts of the commodities have more or less differing prices.   To average these second prices is really to average the deviations of the prices from the common starting price.   These deviations of the prices are nothing but the variations of the prices ; so

[59] Cf. *The Measurement of General Exchange-Value*, pp. 158–9. The process of comparing price-levels is rather like trying, on an always billowy sea, which has a flat bottom, and in which the quantity of water constantly varies, to obtain its levels (such as they would be if it were still) by averaging the various heights of different sections of the waves above the bottom (the zero base). The waves would vary in extent according to the winds, and their averages above the bottom would vary according to changes in the quantity of water.   If the proper averages were correctly drawn, their variations would represent variations in the never-existing level or plane surface of the water, and yet would be just as truthful as if the level did exist.

[60] *A serious Fall in the Value of Gold ascertained*, 1863, reprinted in *Investigations in Currency and Finance*, p. 23.

[61] Cf. *The Measurement of General Exchange-Value*, pp. 165 n., 200–1.   For brevity this work will hereafter be referred to as *M. G. E.-V.*

that, without performing the labour of reducing all commodities to a common price (which would properly have to be repeated every year, or whatever other interval is chosen [62]), the same result is obtained by simply averaging the price-variations.

Again, however, suppose we have performed that labour of reduction, and starting with all prices alike, we average their deviations at the next period. This process must not be likened to the process of averaging observations. The purpose of averaging observations is to get an average observation that is most probably nearest to the true position (a magnitude of some sort) of the object observed. The true position is conceived as, so to speak, the bull's-eye at which the observations have aimed (as the astronomer aims his telescopic instruments at a star) ; but the observations have not exactly hit it, and have deviated, if unbiassed or accidental, almost equally on both sides, and the average of them is believed to bring them back most closely to the true position aimed at. But in averaging price-deviations we do not desire to get back to the starting point ; for that we know already. We desire to get the average deviation. It has some resemblance to the practice indulged in by observers of averaging their errors—that is, their reputed errors, as measured from the average (not from the real and true, but otherwise unknown, position). Astronomers and surveyors perform this, to them secondary, operation only to get the " probable error," as they call it, of their averaged observation, by which they mean the probable amount of accuracy of their observations. All this they do with observations the errors of which are believed to be unbiassed and only accidental. And yet this similitude does not go on all fours, as in calculating the " probable error " the errors on opposite sides are averaged separately. A better analogy is the following.

Suppose a great number of shots are aimed directly at a distant bull's-eye while a strong wind is blowing sideways across the path of the bullets. The wind deflects them all to a more or less equal extent toward the side of the bull's-eye to which the wind is blowing. It thereby extends the deflections of the bullets that are already accidentally deflected on that side, and it lessens the deflections of the bullets that are

---

[62] Otherwise a various and haphazard weighting is introduced for every comparison made between subsequent periods, contrary to the intent of the operator. Cf. *M. G. E.-V.*, pp. 188–9.

already deflected on the opposite side, and may even reduce their deflections to nothing or carry them over to the other side. When, therefore, we average the lateral positions of the hits, the average will no longer approximate to the bull's-eye, but will lie at a greater or less distance to the lee side of it according to the strength of the wind ; and this distance of the average of the hits, or the average hit, from the centre of the bull's-eye, will measure approximately the bias caused by the wind of the given intensity. The same result would be obtained by averaging directly the lateral errors of the hits, this average giving the same bias of the wind in a much shorter process.[63] This operation, in either of its two forms, of measuring the bias that has deflected all the hits more or less alike, is the only one in connection with errors of observation that has complete resemblance to the statistical problem in economics of averaging price-variations, in either of its two corresponding forms. The level of prices or the exchange-value of money remaining the same, prices must vary by both rising and falling equally (but remember there are different ways of reckoning equality) : their rises and falls must compensate one another, or else the price-level and the exchange-value of money have not remained unchanged. If they vary more on the one side than on the other—falling more than they rise, or reversely, rightly reckoned—they may be said to show a bias, which is the change of the price-level, or inversely of the exchange-value of money, caused by whatever is the cause of change in the exchange-value of money (roughly a change in the quantity and mobility of money relatively to the quantity and mobility of commodities). It may be added that riflemen measure the bias caused by various intensities of the wind blowing in various directions, in order to be able to allow for it, after measuring the intensity and direction of the wind when they shoot, by aiming the ascertained angle to the windward side. In economics what corresponds to this is an inverse operation, or one like measuring the deflecting power of the wind by measuring the bias of the hits on the target. For here the causes are more complex even than the effects, so that we

[63] Only approximately correct is the result because it is *plus* or *minus* what would have been the true error of the average from the centre of the bull's-eye, had there been no wind. But this error may be small on the supposition of a great number of fairly accurate shots. In astronomical observations the bias of the wind is replaced by the bias of refraction, according to the height of the star above the horizon.

cannot argue from them to the effects, but we must measure the effects first and directly. The deviation of the whole price-level being thus measured, it may now be allowed for, by altering contracts accordingly ; or an attempt may be made to counteract it—to prevent its increase, and to bring it back to the starting point—by decreasing or increasing the currency as soon as a rise or a fall of the price-level manifests itself. But this is a use to be made of the measurement of variations in the exchange-value of money, supposed to be already accomplished, and lies beyond our subject ; as also does the consideration of the causes of the variations that are measured. The averaging of price-variations in order to measure the inverse variation in the exchange-value of money, is one measurement ; and the measurement, whatever may be its nature or procedure, of the causes of such variations, is another distinct measurement.[64] The former is our present subject.

But properly the similitude of averaging prices is not to what is done in connection with observations, but to what is or should be done in connection with estimates. Only these have not been so thoroughly studied by others ; wherefore it is incumbent on statisticians to study, not so much the averaging of estimates, which is not their affair, but directly the averaging of price-variations.

Still, to the extent that the subject of estimation has been developed, it may be of some use as a model. For price-deviations have many of the peculiarities possessed by the aberrations of estimates. Thus, as it is possible for an estimate to be made as high as any quantity whatever, so it is possible for a price to rise to any quantity whatever ; and as it is impossible for an estimate to be made as low as zero, and certainly not lower, so it is impossible for any price to

[64] It has escaped the attention of many economists that if they are to argue from the causes of price-variations to the price-variations, they would have to be able first to measure those causes quantitatively. But, with the possibly single exception of Professor Irving Fisher, they have not yet even reached the stage of entering upon the subject of this kind of measurement, since they have not yet come to agreement about the nature of those causes. Difficult as is the measurement of price-variations, much more difficult would it be to measure their causes with anything like equal accuracy. We may, therefore, place much greater reliance upon the direct measurement of price-variations, than we could gain by bringing in any reference to their causes. Cf. *M. G. E.-V.*, pp. 22-3.

sink to less than zero and even to zero, for then it would cease to be a price and the article would go out of the market.[65]   Now, when in the case of two articles at first priced the same (for some quantity of one article is always equivalent to some quantity of another) the price of the one rises and the price of the other falls, we have almost exactly the same problem and its modifications that those seventeenth century disputants had with their hypothetical case of estimates higher and lower than a true value.   For we may wish to know whether the rise of the one is greater than the fall of the other ; or, the rise being given, what the fall must be to equal it, and reversely ; or again, the two deviations being given, what is their combined effect upon the price-level, that is, what common variation of the two prices together would have the same effect—which common variation must evidently be a mean of some sort between the two given variations.   These are not idle questions like those debated so violently by Nozzolini and reluctantly by Galileo ; for if the price-level has changed, we want to know it, so as to adapt ourselves to it, and, too, to counteract it, if possible. Thus that academic dispute of nearly three centuries ago has a direct bearing upon a very important problem, with which not only scientific but also practical people are to-day much concerned.

If we confine our attention to only two prices, of two articles of equal importance, almost everything that was advanced in that old controversy would seem, *mutatis mutandis*, to have application.   If one price rises to more than double what it was, it cannot be equalled by any fall of the other price, according to the arithmetic mean, just as, according to the harmonic mean, if the one falls to less than half its previous price, it cannot be equalled by any rise of the other ; but whatever the rise or the fall of the one, it can be equalled by some fall or rise of the other, according to the geometric mean.   Again, if the one price rises from 80 to 100 and the other falls from 100 to 80, these obviously are an equal rise and fall ; but equality is assigned to them only by the

[65] Above, in note 12, it was said that we might suppose a boot to be given with an article and *conceive* this boot as a negative price.   But that supposition does not take place in ordinary commercial transactions.   Jevons noticed that we pay to get rid of some things, and such things he considered to have negative value, *Theory of Political Economy*, 4th ed., p. 127.   But such things do not appear in the market, nor are they consumed. Cf. *M. G. E.-V.*, p. 248.

geometric mean. The question is not one of equal differences or distances, but of equal ratios, equality of difference coinciding with equality of ratio only in cases like the last. No one can read the arguments of Galileo and Castelli without being convinced of their superiority to the arguments of Nozzolini. And the same arguments apply here. They apply here, however, only on the assumption above made—an assumption involved in that other problem itself, but here not easily carried out (we shall return to it later)—that the two deviations, and consequently the two articles, are equally important, so as to permit them to be weighted evenly.[66]

It will be objected that in the case of prices we are not concerned with only two articles, but with a whole marketful of them, and now the contingency exists that any rises and falls of some prices can be balanced by the falls and rises of other prices even according to the arithmetic average (and, it may be added, according to the harmonic, too), since now any one extravagant rise (or fall) can be met by falls (or rises) of several other articles. Yet even so there is a possible case, which would admit of no balancing, as when all the prices happen to come under two groups (or under three, the third being of prices that have not changed), in the one of which the prices rise (or fall) extravagantly, and in the other the prices fall (or rise), but cannot fall (or rise) so far as is required by the arithmetic average (or the harmonic). Apart from this, which must be allowed to be an improbable case, the argument against the arithmetic average (and the harmonic) is that it would be strange if, the geometric mean being the right one to use in the case of two prices, the arithmetic mean or rather average (or the harmonic) suddenly became the right one to use in the case of three or four or more prices—or at what number is the jump to be made ? The dispute we have reviewed over errors of estimation might very readily have been extended to more estimates than two. Can we believe that Galileo, after demonstrating (for this is not too strong a word) the correctness of the geometric mean for measuring the errors of two estimates, would have adopted the arithmetic average for measuring the errors of many estimates ? The law of continuity, which Galileo recognised, forbids.

It would seem as if it could be laid down as axiomatic, that the same mean or average must be applicable to any possible number of items. Two price-deviations are possible items.

<hr>

[66] Cf. *M. G. E.-V.*, pp. 248-9.

Therefore a mean or average right or wrong for two such items, it would seem, must be right or wrong for any other number of them. But now a curious phenomenon shows itself, of which a hint has already appeared. This is, that in the case of the geometric average, although it would seem to be the theoretically right one for any number of price-variations as well as for two, yet it works with perfect accuracy only of two (equally important) price-variations. The axiomatic test, then, can be stated only as a negative one. A mean or average not applicable to the case of two price-variations cannot be the right one for the case of many price-variations—that is, the theoretically right one, although in practice it may approximate to a theoretically right one. The justification for trying a problem of averages on two objects first, as is so often done by statisticians and others (we have seen Bertillon and Bowley doing so), is of course that such are the simplest cases. If the average in question will not work with two objects, it is not the theoretically correct average. But if it will work with two objects, we need to go on and apply it to the more complex cases of many objects. Two objects form a perfect negative test : they may disprove an average. They are not a positive test : they cannot prove an average—or of any kind of average it must first be proved in general that they can. This is provable of the arithmetic and harmonic averages,[67] but not of the geometric average.[68] The geometric average, if in any subject it passes the first ordeal, needs to be examined further in application to many objects.

[67] In a paper on the method of least squares in the *Berliner Astronomisches Jahrbuch*, 1834, Encke tried to prove that method by demonstrating that because the arithmetic mean is correct for reducing two observations, therefore the arithmetic average is correct for reducing three or more observations. A generation later Venn in his *Logic of Chance*, Chapter XIII. § 13, expressed the opinion that, although we must adopt the arithmetic mean for reducing two observations, we have a wider choice of inter-mediate values when we come to reducing three or more. It was probably to allay such doubts as this in his day, that Encke made his demonstration. It needs hardly to be said that he had little difficulty.

[68] This distinction was first pointed out in *M. G. E.-V.*, pp. 239–42, cf. 519–20. That work may be said to be founded on this distinction, since otherwise it would have proved the geometric average, and the subject would have been disposed of in short order. Cf. p. 255.

It is well known that when the problem of measuring variations, or deviations, of prices came up for scientific investigation, the arithmetic average was at first adopted without hesitation and without examination, and continued to be so used till Jevons perceived that it was a case for proportions instead of differences, since variations of prices are changes of ratios, and accordingly recommended the geometric average.[69] It is clear from Jevons's first statements on the subject, that he had a correct idea of the geometric deviation involved. He said that not the arithmetic but the geometric average must be used in his inquiry about the change going on in the value of gold, because the alterations of prices with which he was dealing had a wide range, " varying from more than 50 per cent. of decrease to more than 100 per cent. of increase " ; and he instanced as almost exactly such, cocoa and cloves, the former of which had risen from 100 to 200, and the latter had fallen from 100 to 50, so that, in his words, " the price of one is doubled, of the other halved—one is multiplied by two, one divided by two " ; just as, in the seventeenth-century problem, the true value was multiplied in the one estimate by ten and in the other it was divided by ten ; wherefore Jevons, like Galileo, considered that " on the average " the prices of these articles had remained as they were before,[70] their opposite deviations really being equal and compensating each other. Only Jevons did not state this last proposition about the geometric equality of the opposite deviations, did not generalise and formulate it, did not reach the bottom of the subject. And as far as he did go, he did merely for one set of deviations, but hastily and without demonstration, what Galileo had more fully done for another. We may therefore say, that the first advocate of the geometric mean or average for the measurement of deviations, when they are proportional, was a greater genius even than Jevons. It is another instance, also, of astronomy leading the way for statistics. And it is curious to note a further parallelism in the fact that Jevons's position was disputed by Laspeyres as Galileo's had been disputed by Nozzolini ; and Laspeyres converted the question into one of gain and loss, just as Nozzolini had done. We shall have occasion later to revert to this nineteenth-century controversy.

[69] First in the pamphlet *A serious Fall*, etc., in *Investigations*, p. 23. Later in his *Theory of Political Economy*, 1st ed., pp. 81, 83, he defined exchange-value as a " ratio of exchange."

[70] *Investigations*, pp. 23–4.

Jevons did not confine his recommendation of the geometric average to index-numbers. " In almost all the calculations of statistics and commerce," he wrote, " the geometric mean ought, strictly speaking, to be used." [71] This is an exaggeration, but the amount of truth in it has been too little heeded by statisticians. At most they admit that the geometric average has an advantage in smoothing off extravagant fluctuations, and therefore some of them advise its use when it differs much from the arithmetic. [72] This is, as it were, a concession made by the arithmetic to the geometric average, whereas in truth the concession should be the other way : the arithmetic average may be used where it does not much differ from the geometric. That is, this concession can be made to the arithmetic average, on account of its convenience, in subjects where the geometric is the theoretically correct average ; but where the arithmetic average is the theoretically correct one, no concession on its part need ever be made to the geometric average, for the geometric average, being the more difficult, should never be used except where theory demands it. Thus the procedure of these economists who concede the use of the geometric average for extravagant cases, is really an unconscious tribute to the theoretical correctness of the geometric average. A similarly unconscious recognition of it is made by some economists, who, of late, have advocated the drawing of logarithmic diagrams, and for this purpose the use of logarithmically ruled paper. [73] For, when two or more companion curves are drawn in this way, the eye naturally makes a *résumé* of them by imagining a line halfway between, or even this is actually inserted by the operator ; but the arithmetic mean or average (represented by this line) between logarithms gives the geometric mean or average between the natural figures. [74] Using logarithms also,

[71] *Principles of Science*, 3rd ed., p. 361 ; cf. *Investigations*, p. 128 n.

[72] *E.g.*, Bowley, who calls it a " tentative rule," *Elements of Statistics*, pp. 128–9 ; followed by Zizek, *Statistical Averages*, p. 197.

[73] *E.g.*, Bowley, *op. cit.*, pp. 188–96 ; I. Fisher, *The " Ratio " Chart for plotting Statistics*, Quarterly Publications of the American Statistical Association, June, 1917. The latter's bibliography shows that Jevons (see *Investigations*, p. 128) was the first to use and recommend this kind of charting. It fitted in with his recommendation of the geometric average.

[74] Note that in this connection Bowley says : " For the pur-

Fechner has converted the Gaussian law, originally applied to the arithmetic average, into a law applicable to the geometric average of deviations. He did so for the reason that the deviations on large objects are absolutely greater than the deviations on small objects, although they are not necessarily so proportionally, as, to use an instance he cited, the deviations of the sizes of men compared with those of fleas. For the measurement of all deviations, therefore, including the average stature of men, which we have seen to be correctly averaged arithmetically, he thought the geometric average to be the strictly and theoretically ("prinzipiell streng," strictly according to theory) correct one, although he would concede the use of the arithmetic average when the deviations are comparatively small, since then it is inappreciably different from the geometric.[75] But the reason he gave is not enough. For the use of the geometric average it needs to be shown, not only that the deviations of large objects are greater than those of small ones, but that the deviations above the normal position are or can be greater than the deviations below it. This Fechner did not take into account, and so he applied his method to such objects as the dimensions of visiting cards and of pictures. He merely experimented with the geometric average, without any correct theoretical guidance, and without any satisfactory theoretical or practical results.

Credit for first perceiving the resemblance of prices to estimates and thereby getting an insight into the true nature of their variations, belongs to Profesor Edgeworth. After adopting, in 1887, in the passage already quoted, Galton's re-discovery of the geometric nature of estimation, and referring to Fechner's and Weber's law, he continued : " This law of prizing may well extend to prices " ; and he added : " There exists a simple reason why prices are apt to deviate more in excess than in defect, namely that a price may rise to any amount, but cannot sink below zero."[76] In this he

poses of price index-numbers it is ratios which are important and which the diagram should represent," p. 190.

[75] *Kollektivmasslehre*, pp. 82–3, and so pp. 24–5, 77–82, 91, 95, 307, 339–64. The opinions of this posthumously published work he had reached twenty years before, for their outline is found in his paper *Ueber den Ausgangswerth der kleinsten Abweichungssumme* of 1878, pp. 14–15.

[76] *Memorandum* to the first Report of the British Association Committee, p. 283. He remarked that Venn had just called this kind of rising and falling " a one-ended phenomenon." It may

adduced a better reason than Jevons had given for the geometric average in connection with prices, or rather gave it more explicitly.   But he went on and invoked another. " It appears," he wrote, " that prices group themselves about a mean, not according to a symmetrical curve like that which corresponds to the arithmetic mean, but according to an unsymmetrical curve like that which corresponds to the geometric mean." [77]   And he added some " empirical evidence " in support of this statement.[78]   This last is not so good a reason as the one first quoted, which Professor Edgeworth did not bring into immediate connection with the geometric average.   The primary reason for the propriety of the geometric averaging of price-variations is because prices have a lower and no upper limit to their movements. All that can subsequently be discovered about their multiple dispersion may be interesting, but it is superfluous.   The criterion of dispersion we have seen to be needed only for dubious cases.   This is not a dubious case.   The geometric average is here determined by a perfectly good criterion. To repeat an illustration already used : Professor Edgeworth, we can have no doubt, would not ascribe the reason for our employing the geometric average in averaging the annual rates of increase of a country's population over a number of

be noted that Quetelet had noticed (on very insufficient data) that prices deviate more above than below their mean, *Lettres sur la Théorie des Probabilités*, pp. 72–3 ; but he inferred from this only that the causes at work for raising prices are more energetic than those for lowering them, p. 183.

[77] *Memorandum* to the First Report, *ibid.*  In his *Memorandum* to the second Report, p. 208, he claims that, so far as he knew, he was the first to advance this theory of prices ; and he referred to it again as his discovery in *New Methods of Measuring Variations in General Prices*, in the *Journal of the Royal Statistical Society*, June, 1888, pp. 351, 356.

[78] *Memorandum* to the first Report, pp. 284–6.   Professor Wesley C. Mitchell has likewise given an analysis of over 5,000 price-variations, which supports it, in *Index Numbers of Wholesale Prices*, in the *Bulletin of U.S. Labor Statistics*, No. 173, 1915, pp. 11–24 ; where his table No. 2 presents a specimen of almost perfect geometric dispersion around the median.   But very curiously Professor Mitchell calls it, p. 18 " a distribution according perfectly with the so-called ' normal law of error,' " like errors of observation, etc., as if it exhibited arithmetic dispersion, while on pp. 20–1 he recognises that it is not " perfectly symmetrical " although it is " regular."   On p. 82 he, too, notes that prices can rise without limit, but have a limit to their fall.

years to the fact, if fact it be, that such annual rates group themselves around their average according to an unsymmetrical curve which is the regular curve representative of geometric dispersion. And if the ulterior study of dispersion is not needed in this case, it is not needed in the case of price-variations. Because of not resting at the simple, true, and sufficient reason, Professor Edgeworth has unfortunately been led astray into studies of dispersion involving the calculus of probabilities, which, however interesting in themselves and useful for their own purposes, are not pertinent to the study of averaging price-variations for the purpose of measuring variations in the exchange-value of money ; but which Professor Edgeworth has come to consider basic to this study. There will be need of going into this question more fully later.

## Weighting.

Prices not only have some resemblance to estimates, and through them to observations, but they have some difference. They are properly variations, while estimates and observations are deviations. For this reason their dispersion around a centre is of no importance. The problem, as has been shown, is not to get from them the original centre : it is to get an average of their departures from a common starting point, or more simply, an average of their variations. This average may form a new centre for new departures—or rather a new level for new risings and fallings.

Among the various problems we have reviewed, one having partial analogy with the problem of axiometry is the problem of getting the true position when the estimates and their errors are both known. For here what corresponds to the estimates are the prices, and what corresponds to the errors are the deviations or variations of the prices. Prices and their deviations are here the two factors given, and sought is the value around which the deviations compensate each other and balance, just as in that other problem the position sought is one around which the erroneous estimates, properly weighted, equal each other and balance. But between the present problem and that problem a great difference comes in with the question of weighting. There the weighting is given with the other two elements, since it is inversely according to the errors. Here no such weighting can be employed ; for the variations of the prices have influence on the weights only as they affect another set of

factors, which exist here but have no existence there.    These
other factors are the full values of the articles whose varying
prices are being averaged.    The full value of any kind of
article itself depends upon the price of the article and the
quantity of the article (or the number of its units so priced)
that come under consideration.    This element of quantity,
and the full value which it produces with the price-element,
have nothing analogous to them in the problem of the errors
of estimates or observations ;  and so they very seriously
differentiate the problem of averaging price-variations from
the problem of averaging errors of estimates and observa-
tions—and, it may be added, from the problem of averaging
estimates or observations in any of its forms.

But the principle of weighting is the same at bottom in all
kinds and instances of averaging.    In the words of Professor
Merriman, " weights are merely numbers denoting repeti-
tion." [79]   Certain items are to be averaged :  several of them
may be the same quantity or magnitude ;  then, instead of
treating such equal items as so many distinct items, we
combine them into one item several times repeated, and the
coefficient representing its repetition is its weight.[80]   Thus
in the problem of finding the true position from two known
estimates and their known errors, the estimate with the
smaller error has to be repeated in order that its error may
equal and so compensate the larger error of the other estimate.
In averaging estimates or observations of which the errors
are unknown, in order to obtain the nearest approach to the
true position, the estimates or observations must be treated
as single items (and then if two are alike, they are combined
into one item twice repeated, and so on), provided they have
been made with equal care and therefore can be considered
as equally good.    But if some have been made with greater
care, or with finer instruments, we may consider them as
being twice or thrice as good as the others, and so we may
treat the former as equivalent to two or three of the latter,
and weight them accordingly.[81]   The question of weighting,
then, is a question as to the nature of the single items that
are to be averaged.[82]   In price-variations what are the single
items ?

It might seem at first sight as if simply every price-
quotation were a single item, and since every commodity

[79] *The Method of Least Squares*, § 51, cf. § 39.

[80] Cf. *M. G. E.-V.*, pp. 87–8.

[81] Cf. Venn, *Logic of Chance*, p. 346.

[82] Cf. *M. G. E.-V.*, pp. 89–90.

(any kind of commodity) has one price-quotation attached to it, it would seem as if price-variations of every kind of commodity were the single item in question. This is the way the question struck the first inquirers into price-variations, wherefore they used simple averaging with even weighting. But a price-quotation is the quotation of the price of a generic name for many articles ; and one such generic name covers a few articles, and another covers many. The few and the many articles themselves consist of so many pounds or bushels or other denominations of measures ; but these as such have nothing to do in the economic world of values—that is, they do not themselves appear in that world, though they act, so to speak, as substances underlying it. The generic name of a commodity, covering so many pounds or bushels at a certain price, covers so many value-units of those singly-named articles, and these value-units evidently are the single articles variously repeated. A single price-quotation, therefore, may be the quotation of the price of a hundred, a thousand, or a million dollar's worths, or pound's worths, of the articles that make up the commodity named. Its weight in the averaging, therefore, ought to be according to these money-unit's worths.[83] Or, to reach the same result more briefly : the weights of commodities are directly according to their importance ; and the importance of every commodity is proportional to what may be called its full value, or its quantity multiplied by its price.[84]

This may be rendered still plainer by an example. Suppose wheat is classified into two grades and their prices quoted separately, and barley is quoted only as one commodity. Suppose, further, the full value of wheat is, grade A five millions, grade B five millions, and of barley five millions

[83] Cf. *M. G. E.-V.*, p. 93, cf. pp. 80–1.

[84] A warning may be needed against misuse of the word " importance " in this connection. It is used simply as a convenient term for summing up so many money-units' worth. There is no reference to utility, whether total or final. Utility has no more to do with measuring price-variations than with measuring temperature. Utility may be among the causes of exchange-value ; but knowledge of causes is not necessary for measuring effects. Objective ratios of exchange can be measured without such knowledge, even by observers who have no notion of utility. Nor is there any reference to the fact that a more important commodity " absorbs more currency." It refers only to the fact that in a more important commodity there is a greater number of price-variations. Cf. *M. G. E.-V.*, pp. 8–9, 94.

(money units of any sort being understood).   Evidently we have three equal items.   Now suppose the distinction between the two grades of wheat had not been made and the prices of wheat had all been thrown together.   Then we should have had, wheat ten millions, barley five millions. Evidently the difference between the two cases is only nominal : two items have been merged into one, with only one price-quotation, giving the appearance of one item ; but in truth we still have two items of wheat, or three items all told.   Nor is the matter affected by the fact, or not, whether wheat admits of being classified into two grades.   Let all wheat be alike : yet five millions of wheat is a single item over against five millions of barley, and another five millions of wheat is either another item, or it is the same item repeated.   Therefore we should weight wheat as two compared with barley as one.   For weights are proportional figures, and we naturally reduce them to their least common denominator.

When the early investigators simply took each quotation of a price-variation as one item, we may suspect they were influenced also by the consideration of convenience.   There is prevalent a stupid distinction between weighted and unweighted or simple averages, as if the former contained weights and the latter did not.   The latter contain weights just as much as the former, only they are equal or even weights.[85]   You cannot avoid the subject of weighting by adopting the simple average, for it has pursued you there. You must have either a right or wrong system of weighting, or rather a better or a worse.   If you make an earnest attempt to adopt a good system, something can generally be said for it ; but if you adopt even weighting, nothing can be said for it—except the lazy man's recommendation of superior easiness.   It is, however, argued : all systems of weighting give nearly the same results in practice ; therefore adopt the easiest, which is even weighting.   It is like the argument also made for the arithmetic average compared with the others.   All depends on the word "nearly."   If index-numbers are ever to be put to any serious use, the closest possible approximation to the truth will be necessary. For it will be necessary to make the measurements from year to year, or even oftener, and to combine the results in a series ; and if a method be adopted that gives a small error always or mostly on the same side of the truth, its errors will

[85] Cf. *M. G. E.-V.*, p. 82.

accumulate from period to period.   If, to avoid this danger,
the ordinary system be kept in use of comparing every
subsequent period with the same period as a base, this
system with a faulty average or weighting, not excepting
even weighting, produces a varying haphazard weighting,
with variously resulting errors, in all comparisons between
the subsequent periods, increasing in irregularity as the
series advances,[86]—and it is principally for comparison with
the immediately preceding period that these measurements
are desired.[87]

Besides, we need first of all to examine the matter *in
theory*, by which practice may afterwards be regulated.
Theory demands precision, and can make no compromise
with sloth.   " A life of action," says Macaulay, " if it is to
be useful, must be a life of compromise.   But speculation
admits of no compromise." [88]   Now, in theory the considera-
tion of weighting is absolutely necessary.   And it must be
treated in close connection with the averages that are being
investigated.[89]   For subconsciously various kinds of weight-
ing obtrude themselves in company with various kinds of
averages.   The subject may be illustrated by the controversy
between Jevons and Laspeyres.

Jevons asserted that because the price of one article had
risen 100 per cent. and the price of another had fallen 50 per
cent., it would be " totally erroneous " to say the average
variation of the two prices was a rise of 25 per cent., since
the geometric mean between them indicates no variation
at all, and the geometric mean of the ratios is the right one
to use.[90]   To this Laspeyres objected that as it later requires
200 money-units to purchase the one article which 100 would
have purchased before, and fifty to purchase another which
100 would have purchased before, it now requires 250 to
purchase what 200 would have purchased before, which is
the same as requiring 125 to purchase what 100 would have
purchased before, wherefore the prices of these two articles
together have risen from 100 to 125, as indicated by the
arithmetic mean, or by 25 per cent.[91]   Obvious is the

[86] Cf. *M. G. E.-V.*, pp. 189, 206–7.

[87] Cf. Mitchell, *Index-Numbers of Wholesale Prices*, pp. 36–7 ;
also Fisher, *Purchasing Power of Money*, pp. 203, 423–4.

[88] *Essay on Mahon's History.*

[89] Cf. *M. G. E.-V.*, p. 275.

[90] *Investigations*, pp. 23–4.

[91] *Hamburger Waarenpreise*, 1850–1863, in *Jahrbücher für
Nationaloekonomie und Statistik*, 1864, p. 97.

resemblance of Laspeyres's procedure to Nozzolini's, in immediately converting the problem into one of gain and loss ; while Jevons, like Galileo, treated it as a question of ratios. But Jevons for a moment went over into Laspeyres's camp to dispute with him on his own ground.  He urged that we could with equal reason " suppose a certain uniform quantity of gold money to be expended in equal portions in the purchase of certain commodities, and that we ought to take the average quantity purchased each year " ; which " might be ascertained by taking the harmonic mean " of the prices.[92] This suggestion was actually adopted by Messedaglia.[93] Professor N. G. Pierson reviewed this debate, and because of the inconsistencies it involved, various results being one as good as another, concluded that the whole subject is insoluble.[94]  Now, originally Jevons had no right, as Galileo had had, to use the simple geometric mean, unless he added the assumption that the two articles were equally important, and unless he gave the proper definition of what is to be meant by their equal importance.  Laspeyres, on his part, did make a further assumption, namely that the same quantities of the two articles are purchased at the two periods, notwithstanding the changes in their prices.  This assumption he had a perfect right to make, as such ;  only he should have drawn attention to it, and have noted its arbitrariness, which permitted it to be countered by Jevons's later equally arbitrary assumption of another sort.  Thus Laspeyres's problem was different from Jevons's original problem, because Jevons had not made the assumption which Laspeyres made ;  and it was different from Jevons's later suggested problem, because that rested on another different assumption.  They argued at cross purposes ;  and they did so because they did not introduce the subject of weighting except in this secondary clandestine manner, and they never argued as to which kind of weighting is the more correct.   Galileo, we have seen, advanced some incomplete suppositions, and then remarked that no solution could be made from them, because some essential data were not supplied.   Jevons set before himself a problem with just such insufficient data, and reasoned about them just as if he had given himself all the data needed ; and when

[92] *Investigations*, pp. 120–1.

[93] *Il Calcolo dei Valori Medii*, p. 39.

[94] *Further Considerations on Index-Numbers, Economic Journal*, March, 1896, pp. 129–30.

Laspeyres supplied some data, he supplied some others, and neither examined what the additional data ought to be in order to represent the ordinary state of things. The two commodities under consideration were really two classes containing different numbers of individuals, whose prices, themselves averaged at each period, varied between the periods ; so that the price-variations they were averaging were the price-variations, not of *two* commodities, but of an unknown (or unposited) number of individuals distributed over two classes. Yet neither of the disputants (nor Messedaglia and Pierson after them) made the slightest attempt to ascertain or to define what and how many were those individual things whose price-variations they were averaging. It was as if two anthropologists should hear that the average height of one group of men were so and so, and the average height of another group of men were so and so, and then, without trying to find out how many individuals were in each group, should dispute over the proper kind of average for averaging these averaged heights. It is pitiable to think that men of scientific attainment should argue in this way.[95] And yet others have continued to take sides in this dispute in the same old incomplete way ; and his opponents have criticised Jevons rather for using the geometric average than for neglecting to inquire into the proper weighting. In all his writings only two passages can be found in which, in this connection, Jevons referred to weighting, and then he showed indifference to the subject. He even entertained the idea of weighting his commodities inversely according to their price-variations ![96] In his practical work he was bothered by the high prices of cotton in 1862, and although the geometric average was recommended as especially suitable for such extravagant cases, he actually doctored those prices. Had he noticed that the quantity of cotton had been even more extremely reduced, those high prices, properly weighted, would have given him no cause of uneasiness. It may be remarked that Laspeyres added at the end of his own argument : " What is true of two commodities, is true also of any

[95] Cf. *M. G. E.-V.*, pp. 266–8, 275.

[96] *Investigations*, pp. 21, 57. Jevons was one of the first, if not the first, to mention the use of "ponderation " in the subject of axiometry ; and therefore he may well be excused for not being precise with regard to it, for making a wild suggestion, and for refusing to " attempt to decide " the questions involved. But *we* should advance beyond his first step.

number of commodities." He happened to be right ; for this is true of the arithmetic average with which he was dealing, as also of the harmonic. But neither he nor Jevons perceived that this statement, as it stands in all its generalness, is not true ; for it is not true of the geometric average. Jevons never saw, what we shall soon see, that he had no right to extend to the geometric average of many pricevariations, as he tacitly did, what he might have proved of the geometric mean in the case of two commodities ; and Laspeyres never took him to task for this oversight.

In measuring variations of the exchange-value of money we pass from one period to another. Now, at any one period the importance of every commodity is measured by its full value, which is calculated by the quantity of the commodity that has been dealt with during that period multiplied by its price. If we are measuring the variation of the money of a country (not of any one class of its population), we should in strict theory take into account all commodities (along with some other things, possibly, such as house-rent, but not wages), and the complete quantities of them that have appeared in the markets throughout the whole country. This is evidently impossible, and samples of the most simple and useful commodities have to be selected, and their quantities have to be estimated more or less accurately. Some nice questions arise as to whether only what is consumed in the country, or only what is produced in it, or both together are to be counted ; and also there are difficulties as to the single price-quotation that is to be given at each period to each commodity, since this, too, must be an average.[97] If,

[97] Throughout the country during the period a commodity is not sold at one price, nor even at one wholesale price in its principal market. Various quantities of it are sold at different prices, and the full value is obtained by adding all the sums spent (at the same stage in its advance towards the consumer), and the average price is found by dividing the total sum (or the full value) by the total quantities. This result is the arithmetic average of the prices weighted according to the quantities sold at the prices. Between producer and consumer some goods pass through few and some through many hands. It would seem as if only one sale (say the first or the last) ought to be counted ; but this has been disputed. As it is impossible to record all sales, estimates have to be relied on ; which leaves room for error, and for the calculation of error by the doctrine of probabilities. This doctrine shows that, when the investigation has been conducted with care and without bias, because of the likelihood of the errors

further, instead of its being the money of the country that is being measured, it is the money of some class (say labourers), then a different selection of articles, with different quantities, and even with different prices (retail instead of wholesale), has to be made.    Into these details it is impossible here to enter, and it is not necessary ; for they have no effect upon the problem of averaging price-variations, which is the problem before us.    This problem presupposes that the prices and the quantities are given,[98] and it then proceeds to operate upon them.    This problem is a large one in itself, and quite deserving of being isolated from the problems involved in the providing of its data.    It is a mathematical problem, needing to be solved in theory first, while those are mostly practical questions, for practical men to decide.    The theoretician may leave those problems to others.    The others should take the solution of the problem of averaging from the theoretician, modifying it to suit their practical purposes.    Being utterly distinct, this problem and the other problems should never be mixed.

We then assume that (1) commodities, (2) their quantities, and (3) their prices are given, wherewith (4) their full values are also given, whereby their relative importance is measured and their weights in the averaging determined.    The commodities, as classes of things, must be the same at both periods. If a commodity exists at one period and not at another, it must be omitted from any comparison of these two periods ; for its prices have not varied, it has merely appeared or disappeared.[99]    The quantities, however, as well as the prices, are apt to vary between the two periods, involving also variations in the full values.    But we are not averaging the variations of the quantities, nor of the full values : we are averaging only the variations of the prices.    Yet the variations of the quantities, along with the variations of the prices, by affecting the full values cause variations in these ; and the variations of the full values cause variations in the weights wherewith the price-variations are to be averaged.

being on opposite sides and therefore neutralising one another, the effect of the individual errors upon the final result is likely to be small.

[98] This is the position taken by Professor Edgeworth at the beginning of his *Memorandum* to the first Report of the British Association Committee, where, p. 254, he wrote : " It is supposed that the prices of commodities, and also the quantities purchased, at two epochs, are given."

[99] Cf. *M. G. E.-V.*, pp. 112–13.

Thus, after all, the variations of a commodity's price, in so far as it affects the commodity's full value, affects the weight to be assigned to that commodity's price-variation, and so restores a partial resemblance to the weights in the problem of estimates above cited, where the weights are entirely determined by the deviations, or errors, of the estimates, but inversely.   Here, however, the influence, as far as it goes, is neither inverse nor direct, being itself affected by another factor.   For here there is another factor, the quantities, which is entirely absent from the other problem.

With respect to quantities, the present problem has more analogy with the problem of estimates set up by Dr. Bowley. For in Dr. Bowley's problem estimates of two or more different things were involved, and the quantities, or sizes, of these things had influence.   But here still another difference subsists.

Commodities are to be weighted according to their importance, or their full values.   But the problem of axiometry always involves at least two periods.   There is a first period, and there is a second period which is compared with it. Price-variations have taken place between the two, and these are to be averaged to get the amount of their variation as a whole.   But the weights of the commodities at the second period are apt to be different from their weights at the first period.   Which weights, then, are the right ones—those of the first period ? or those of the second ? or should there be a combination of the two sets ?  There is no reason for preferring either the first or the second.   Then the combination of both would seem to be the proper answer.   And this combination itself involves an averaging of the weights of the two periods.

A hint of this difficulty appeared in Dr. Bowley's problem. And there we saw that a wrong system of weighting with a wrong average may give the right result.   Possibly the same may take place here.   Perhaps, then, a wrong average with a wrong weighting may serve us, even though the right average with the right weighting may fail us.   It is all the plainer how close is the connection between the systems of weighting and the averages.[100]

[100] In their dispute above reviewed, Jevons simply treated the two commodities as equally important.   Laspeyres posited that they were equally important at the first period only.   Then Jevons in his reply posited that they were equally important at both periods.   In this way they were able to use different evenly

## Index-numbers.

In averaging the price-variations of varying quantities and full values of commodities, the whole ground is covered by three possible suppositions.    (I.) The quantities may be the same in both periods, and the full values may vary with the prices.    (II.) In spite of variations of prices, the full values may be the same in both periods, because of inverse variations of the quantities.    (III.) Both the prices and the quantities may vary irregularly, and therefore also the full values vary.    The first two suppositions are very improbable, as the third is the one which almost always takes place ;  for the usual course of things is that, the exchange-value of money remaining nearly stable, in most commodities a rise of price generally goes with a diminution of quantity, and a fall of price with an augmentation of quantity, or reversely (for here, that is, for our purpose, it is indifferent which is cause and which effect).    These opposite movements may tend to keep the full values more stable than either of the factors. But there always is much irregularity.    Now, the first two suppositions contain separately the elements into which the third can be analysed ;  and therefore, as logical method enjoins going from the simple to the complex, we may briefly review the matter in this order, since the first two suppositions, though not probable in practice, are possible in theory.

Throughout the following let us represent quantities by $q$, prices by $p$, and price-levels by $P$.    The quantities and prices of particular commodities may be distinguished by accents, and those of different periods by subscript numerals. Full values, then, which may be represented by $v$, are $= qp$.

(I.) The quantities being the same in both periods, there is no need of subscripts for them.    The case is formulated thus :

$$\frac{P_2}{P_1} = \frac{q'p_2' + q''p_2'' + q'''p_2''' + \ldots \text{ to } n \text{ terms}}{q'p_1' + q''p_1'' + q'''p_1''' + \ldots \text{ to } n \text{ terms}}$$
$$= \frac{\Sigma q p_2}{\Sigma q p_1} = \frac{\Sigma v_2}{\Sigma v_1}. \tag{1}$$

weighted averages.  But neither of them inquired whether the weighting they chose fitted the average they used.  They may have hit upon a right combination, or not, but they put forth no effort to prove what was right or wrong in the matter.  They did not posit all the possible, nor even the most likely, combinations of weighting.  They refrained from any systematic inquiry into the subject.

It is evident that the price-level has varied according to the sums of the full values (and the exchange-value of money has varied inversely). A whole people (or a sample portion of it) has purchased the same quantities at different prices. This is the same as saying the average person has bought his portions of the goods, the same in both periods, at different prices. Obviously if he has bought them all with the same expenditure, their average price has not changed ; if he has bought them with a greater expenditure, their average price has risen ; and if he has bought them with a lesser expenditure, their average price has fallen ; and the rise or fall is proportional to the greater or lesser expenditure.[101]  The case here is one of gain and loss.  Suppose the total expenditures are the same ; this means that what you gain by the fall of one price you lose by the equal rise (arithmetically reckoned) of another, or by a distribution of it over several others.  Or again, if the price of one article has risen 50 per cent. (in the usual way of reckoning percentage), it can be compensated by another falling 50 per cent., or by two others falling 25 per cent. each, or the one 30 per cent. and the other 20 per cent., etc.  It is Nozzolini's interpretation of the problem of estimation over again.  But how about the average of the price-variations ? Is this necessarily and only the arithmetic average, as that example would seem to require ? No ; for the weighting that has to be employed in the averaging brings in diversity.

The equation of formula $b$ above (p. 31) exactly to two other formulæ and approximately to a third, can be imitated here.  The formula just given is equal to this,

$$\frac{P_2}{P_1} = \frac{\dfrac{p_2{'}}{p_1} \times q'p_1{'} + \dfrac{p_2{''}}{p_1{''}} \times q''p_1{''} + \dfrac{p_2{'''}}{p_1{'''}} \times q'''p_1{'''} + \cdots}{q'p_1{'} + q''p_1{''} + q'''p_1{'''} + \cdots}$$

$$= \frac{\Sigma v_1 \dfrac{p_2}{p_1}}{\Sigma v_1} \tag{2}$$

which is the formula for the *arithmetic average* of the price-variations with weighting according to the full values of the *first period*.  It is also equal to this,

---

[101] If the prices all vary alike, say by $r$, every $p_2$ would be $rp_1$, and every $v_2$ would be $rv_1$, and the variation of the price-level would be $r$.  If the quantities all vary alike, say by $s$, every $q_2$ would be $sq_1$, and the variation of the price-level could be obtained by dividing the result by $s$.

$$\frac{P_2}{P_1} = \frac{q'p_2' + q''p_2'' + q'''p_2''' + \cdots}{\frac{p'_1}{p_2'} \times q'p_2' + \frac{p_1''}{p_2''} \times q''p_2'' + \frac{p_1'''}{p_2'''} \times q'''p_2''' + \cdots}$$

$$= \frac{\Sigma v_2}{\Sigma v_2 \dfrac{p_1}{p_2}}, \tag{3}$$

which is the formula for the *harmonic average* of the price-variations with weighting according to the full values of the *second period*.[102] It is also found by trial to be, mostly, approximately equal to this,

$$\frac{P_2}{P_1} = \sqrt[\Sigma q \sqrt{p_1 p_2}]{\left(\frac{p_2'}{p_1'}\right)^{q' \sqrt{p'_1 p'_2}} \cdot \left(\frac{p_2''}{p_1''}\right)^{q'' \sqrt{p_1'' p_2''}} \cdot \left(\frac{p_2'''}{p_1'''}\right)^{q''' \sqrt{p_1''' p_2'''}} \cdots}, \tag{4}$$

which is the formula for the *geometric average* of the price-variations with weighting according to the full values over *both the periods*.[103] Thus Galileo's advocacy of the geometric average fares here almost as well as Nozzolini's advocacy of the arithmetic. For Nozzolini could advocate the arithmetic average here only with a weighting for which no reason can be given, and which is no better than another weighting which, used with the harmonic average, gives the same result, so that he must share the honours of the arithmetic average with the harmonic average ; whereas the geometric average, itself apparently the right one, is used with the right weighting, so that everything about it seems to be in its favour. Yet it fails, since it does not give with perfect accuracy the result known to be the right one. We have the paradox, that wrong averages with wrong weightings give the right result, while the right average with the right weighting does not.

[102] If the $p_1$'s are all the same, formula 2 reduces to $\dfrac{\Sigma v_1 p_2}{p_1 \Sigma v_1}$.
If the $p_2$'s are all the same, formula 3 reduces to $\dfrac{p_2 \Sigma v_2}{\Sigma v_2 p_1}$.

[103] The full values of the two periods are compared, not by their arithmetic, but by their geometric mean. The reason for this is given in *M. G. E.-V.*, pp. 105–10, 118–21, 387–8. For the proof of this formula's approximation to the others, see above, note 25. Further cf. *M. G. E.-V.*, p. 350. The first two identifications are repeated by Professor Fisher in *The Purchasing Power of Money*, pp. 365, 397, cf. 393–5 ; but he apparently has no use for the similitude of the third additional formula (4). And the first alone is repeated by Professor Mitchell, *Index-Numbers of Wholesale Prices*, pp. 92–3 and n.

The geometric average, however, does give the right result with perfect accuracy in one still more narrowed case, when, in fact, it is reduced to being the geometric mean between two equally important and hence evenly weighted price-variations. If we deal with only two commodities, or with two groups of commodities, and if their full values over the two periods together are equal (geometrically measured, that is by their products), then it is demonstrable that the geometric mean of their two price-variations with even weighting is the same as the arithmetic average of them with the weighting of the first period, and as the harmonic average of them with the weighting of the second period. This may be represented simply as follows,

$$\sqrt{\frac{p_2{'}}{p_1{'}} \cdot \frac{p_2{''}}{p_1{''}}} = \frac{q'p_2{'} + q''p_2{''}}{q'p_1{'} + q''p_1{''}},$$

on condition that $q'p_1{'} \cdot q'p_2{'} = q''p_1{''} \cdot q''p_2{''}$.[103a] And so, after all, the geometric *mean* with the right weighting, where it is applicable, does give the right result. It is only the geometric *average* that is at fault.

The reason for this agreement in the case of the geometric mean, depends on the well-known proposition that between only two terms the geometric mean is the geometric mean between their arithmetic and their harmonic means. And the reason why the geometric *average*, in formula (4), applied to more than two equally important price-variations (and even to two that are not equally important [104]), is only

[103a.] The proof is given in *M. G. E.-V.*, pp. 519–20, cf. 351–2. In this case the arithmetic and harmonic means are not subject to the objections brought against them in the case of estimations. Those objections apply to these averages only when used with the weighting proper to the geometric average. When used with their own proper weightings, they give the right result.

[104] For then we no longer have two single items, since one of the named items is composed of more than one real item. If the two articles are only slightly divergent in importance over the two periods (the prices diverging but slightly), the geometric average with its proper weighting will still be very close to the two others with their proper weighting ; but it will diverge considerably from them if the one article becomes three or four times more important than the other, and its discrepancy may even be very great with a still greater divergence of the relative weights : cf. *M. G. E.-V.*, pp. 365–7. Something analogous to this is stated by Professor Fisher, *Purchasing Power of Money*, pp. 412–13 n., although what he there says is applied to the geometric average wrongly, because evenly, weighted.

approximately true (in ordinary cases), depends on the fact that between more than two terms the geometric average is not exactly (except for special reasons), but only approximately, the geometric mean between their arithmetic and harmonic averages, as any one may see who will take the trouble to look into the matter, although this fact is often ignored.[105] And any one can find that the geometric average of all the terms is sometimes above and sometimes below the geometric mean of the other two averages of them ; and the same he will find to hold of formula (4) compared with either of the three preceding formulæ. And we can find this general rule, that if the prices of the preponderating commodities (those with the greater weights) rise above the average price-variation, formula (4) gives a result above the other three, and if they fall below that average, it gives a result below the others.[106] Even so, that is, if it happens (which is not likely) that the prices of all the preponderating commodities vary in the same direction and the prices of all the less important commodities vary in the opposite direction (compared with the variation of the price-level), the discrepancy of formula (4) from the others is found to be small, unless the opposite price-variations are excessive ; and of course in the more usual cases of a mixture of the variations and the weights, such as occur in practice, some of the preponderating commodities rising and others falling in price, and some of the less important commodities falling and others rising, and none or few varying greatly, the discrepancy will be still smaller, and may even be reduced to nothing. And in a series of measurements extending over many periods, as the errors will sometimes be in excess and sometimes in defect, there will not be cumulation, but there will be neutralisation, of the errors, wherefore it is probable that the total error will increase in amount very slowly.

[105] It has been ignored even after the proof of it was first given in *M. G. E.-V.*, p. 517. Thus as late as 1913 Zizek and his American translator assert that " of the three means of the same set of values [no limitation being made to two values] the geometric is always the geometric mean of the other two," the arithmetic and the harmonic, *Statistical Averages*, p. 132 n., and again pp. 196-7 n. In the last passage Zizek illustrates the proposition by an instance of two figures. Had he tried it on three figures, he would have seen his mistake.

[106] *M. G. E.-V.*, p. 365. Unfortunately the rule, there given in italics, by a double slip inverts the true statement. Cf. pp. 321, 521.

(II.) The full values remaining the same at both periods, this supposition means that the community has the same total income in both the periods, and lays it out in the same sums to purchase the various commodities, whatever be the changes of their prices.[107]   If, then, the price of a commodity falls, the community buys more of it ; if the price of a commodity rises, the community buys less of it—and so does the average person in the community.   Hence we can examine the purchasing power of money directly by the total quantities which the same total incomes of the community can in this way purchase in the two periods.   But a complication enters here, in that the comparison cannot be merely according to the quantities purchasable, since these have different preciousness, but it must be according to the exchange-values of these quantities ;   and it cannot be according to the exchange-values of these quantities at either of the periods alone, but it must be according to their exchange-values over both the periods together (of course geometrically reckoned).[108]   Therefore the ordinary units of quantity will not serve, and we must construct units, specially for each comparison, that have the same exchange-value over both the periods.   When this is done, we can measure variations in the purchasing power of money by averaging the quantities purchasable of such units in very much the same way as before, getting three exactly equal formulæ and another approximately equal to the three. But here things are inverted, and the arithmetic average of the quantity-variations must use weighting according to the units purchased in the second period, and the harmonic average must use weighting according to the units purchased in the first period ; but of course the geometric average will continue to use weighting according to the geometric mean of the units purchased in both the periods.   The formulæ thus directly giving the variation of the purchasing power (or exchange-value) of money, can now easily be turned back into formulæ of the usual sort giving the variation of the price-level.   The first of the three equal formulæ (for the arithmetic and harmonic equivalents may be omitted) has either of these two equivalent forms,

[107] If the total income increases or decreases, and the sums spent on every commodity are varied in the same proportion, the common variation may be divided out, leaving the same results as in the text.

[108] For the principle see *M. G. E.-V.*, p. 304 ; for the method and for the formulæ that follow in the text here, see pp. 305–10.

$$\frac{P_2}{P_1} = \frac{q_1'\sqrt{p_1'p_2'} + q_1''\sqrt{p_1''p_2''} + q_1'''\sqrt{p_1'''p_2'''} + \cdots}{q_2'\sqrt{p_1'p_2'} + q_2''\sqrt{p_1''p_2''} + q_2'''\sqrt{p_1'p_2'} + \cdots}$$

$$= \frac{\Sigma q_1\sqrt{p_1 p_2}}{\Sigma q_2\sqrt{p_1 p_2}}, \tag{5}$$

or

$$\frac{P_2}{P_1} = \frac{p_2'\sqrt{q_1'q_2'} + p_2''\sqrt{q_1''q_2''} + p_2'''\sqrt{q_1'''q_2'''} + \cdots}{p_1'\sqrt{q_1'q_2'} + p_1''\sqrt{q_1''q_2''} + p_1'''\sqrt{q_1'''q_2'''} + \cdots}$$

$$= \frac{\Sigma \, p_2\sqrt{q_1 q_2}}{\Sigma \, p_1\sqrt{q_1 q_2}}. \tag{6}$$

And the approximately correct geometric average has this formula,

$$\frac{P_2}{\overline{P_1}} = \sqrt[\Sigma\sqrt{v_1 v_2}]{\left(\frac{p_2'}{p_1'}\right)^{\sqrt{v_1'v_2'}} \cdot \left(\frac{p_2''}{p_1''}\right)^{\sqrt{v_1''v_2''}} \cdot \left(\frac{p_2'''}{p_1'''}\right)^{\sqrt{v_1'''v_2'''}} \cdots} \tag{7}$$

This geometric average errs here just as the corresponding geometric average, in formula (4), errs in the preceding supposition, no more and no less, not cumulatively, but neutralisingly.[109]    And, again applied to two commodities equally important over both the periods, so that the simple geometric mean of their price-variations $\left(\sqrt{\frac{p_2'}{p_1'} \cdot \frac{p_2''}{p_1''}}\right)$ can be used, this method gives the exact truth.[110]

But unfortunately, except in the last narrow case of only two equally important commodities (or two such groups of commodities with similar price-variations in each group), formulæ (5) and (6) are open to an objection to which formula (1) is not exposed—an objection which shows that they are not perfectly true, as that formula is.    It would seem to be evident that if prices or quantities or both vary over a number of periods and then return to exactly the same state at a later period as they were in at the first (being the same quantities and the same prices of every commodity), then the measurement of the variations of the price-level from period to period, when combined and strung out in a series, should at the last period indicate the same level as at the first.[111]    This is only a negative test; for many different erroneous methods—in fact, the very worst method of all,

[109] *M. G. E.-V.*, pp. 321, 323.

[110] *M. G. E.-V.*, p. 330.

[111] *E.g.*, if the comparison of the second period with the first showed a rise of 20 per cent., and the comparison of the third with the second showed a fall of 5 per cent., this is a fall from

Dutot's—satisfy it [112]; but if a method does not satisfy this
test, there is something wrong about it.    This test is precisely
analogous to the test which surveyors use in their triangu-
lations ; for they measure round many lines and angles until
they come back to their starting point, and the two ends of
their measurements ought to fit, and when they do not (as
they never do), they apply the method of least squares to
adjust the error over all the intervening measurements.[113]
So here, after three or more measurements of price-variations
over three or more periods after the starting period, if we
assume as the next period the conditions of the first period,
this last measurement, when combined with the others in a
series, ought to indicate no variation, and if it does indicate
some variation, it has in so many measurements committed
error to that extent.    This test, which because of its return-
ing to the starting point may be called *the circular test*, not
only indicates that a method is erroneous (unfortunately it
cannot indicate that it is correct), but also indicates the
amount of its error, and its direction ; and when various
methods are tried by it in application to the same data, it
indicates fairly well their comparative erroneousness.    It is
not serviceable for testing all methods ; for those which
satisfy it must be tested in other ways.    But it is especially
useful for testing the methods which have otherwise proved
themselves nearly correct.[114]

Now, tried by this test, formula (1), with its two variants,
shows no error at any later period, however extended be the
series ;    and, being otherwise demonstrated, it may be
regarded as absolutely true, on its purely theoretical condi-
tion, extended of course to require constancy of the quan-
tities over all the periods.    But on the condition laid down
in the present second supposition likewise extended to all the

120 of $\frac{6}{120}$, so that the series is 100, 120, 114.    To form the series
100, 120, 114, 100, the comparison of the fourth (= the first) with
the third would have to show a fall of 12·3 per cent.    This, of
course, is the inverse of comparing the third with the first, which
should show a rise of 14 per cent.

[112] *M. G. E.-V.*, pp. 203, 205.    This test is a particular appli-
cation of Professor Westergaard's test.    The negative nature of
Professor Westergaard's test has not always been recognised.

[113] This last operation suggests a hint for axiometry.

[114] This circular test is frequently used for this purpose in
*M. G. E.-V.*, pp. 324–41, 368–70, 389–92, 416–18, and especially
423–33.    When the quantities are not supposed the same at the
final period, a complication arises : see pp. 399–402.

periods, the duplicate formulæ (5) and (6) satisfy this test only in the case of three periods, but not always in the case of four and more periods. The reason is that, in the measurement from the second to the third period the units have been changed, and if the next measurement is a comparison with the initial state of things, it, too, has another set of units. Only if we construct units that are equivalent over the whole series of periods compared, will this formula, adapted to such an adjustment, satisfy this test, and then only for that set of periods ; wherefore it does not admit of extension into the future.[115]  But, again, it is proved by trial, with the use of this test, that the aberration of these equivalent formulæ is even smaller than that of formula (7), being for ordinary price-variations and quantity-variations, such as are likely to occur in experience, very small, and sometimes above and sometimes below, wherefore their errors are not cumulative, but tend to neutralise one another.[116]

All the formulæ so far given, whatever their merits and defects under the conditions supposed, have no practical uses, since the suppositions for which they are invented never occur actually. But, being simpler, they contribute to elucidate the remaining more complex cases, which alone represent the actual economic world. But before going on, it should be emphasised that all that has been advanced, except as to the amounts of error determined by trial, is demonstrative, and especially that in the first supposition formula (1) and its variants, and even formula (4) when restricted to its narrowed condition, are demonstratively true—as true as anything in Euclid. The problem of measuring the variation of the price-level, and inversely of the exchange-value of money, is solved, as far as an economic world represented by that supposition is concerned.[117]

(III.) In the actual world both prices and quantities vary, so that it contains features from both the preceding imaginary economic worlds. Now, neither the arithmetic nor the harmonic average of price-variations as obtained in the first supposition is transferable to this third case, because if we use the former with the weighting of the first period and the latter with the weighting of the second period, we no longer get identical results ; and of course it would be absurd to com-

[115] *M. G. E.-V.*, pp. 334–6 ; cf. 398–9.

[116] *M. G. E.-V.*, p. 341.

[117] *M. G. E.-V.*, pp. 401–2. That, on this supposition, formula (1) is the correct method, has been recognised by many writers on the subject. References are given, p. 544.

pare merely the sums of the full values of the two periods.[118] But the geometric average is transferable from those suppositions to this case, because it requires the use of the weighting of both periods.    The form of it to be used is that given in formula (7); for in formula (4) the $q$ really stands for $\sqrt{q_1 q_2}$, and now $\sqrt{q_1 q_2 p_1 p_2}$ is $= \sqrt{v_1 v_2}$, and formula (7) is merely a more comprehensive, or distinguishing form of formula (4).    Now, if our third supposition is confined to a world with only two commodities, or of two groups of commodities in each of which the prices vary alike, and if, further, these two have the same importance over both the periods compared, then the simple geometric mean of these price-variations yields the absolutely true measurement of the variation of the price-level.[119]    But our misfortune is that the geometric average applied to more than two equally

[118] Suppose we draw the arithmetic average of the price-variations with the weighting of the first period: this gives us

$$\frac{P_2}{P_1} = \frac{\Sigma\left(\frac{p_2}{p_1}\right)p_1 q_1}{\Sigma p_1 q_1} = \frac{\Sigma p_2 q_1}{\Sigma p_1 q_1}, \text{ which is the method recommended}$$

by Laspeyres.    Or suppose we use it with the weighting of the

second period: thus $\dfrac{P_2}{P_1} = \dfrac{\Sigma\left(\dfrac{p_2}{p_1}\right)p_2 q_2}{\Sigma p_2 q_2} = \dfrac{\Sigma\left(\dfrac{p_2{}^2}{p_1}\right)q_2}{\Sigma v_2}$, which is a

method recommended by Palgrave.    An arithmetic mean between these has never been recommended.    Evidently none of these have any validity.    Nor would the harmonic average with any such weightings.    Yet the harmonic average with the

weighting of the second period gives $\dfrac{P_2}{P_1} = \dfrac{\Sigma p_2 q_2}{\Sigma p_1 q_2}$, which has been

recommended by Paasche; and with the weighting of the first

period it gives $\dfrac{P_2}{P_1} = \dfrac{\Sigma v_1}{\Sigma\left(\dfrac{p_1{}^2}{p_2}\right)q_1}$, which has been recommended by

nobody.

[119] As do also formulæ 6 (already given) and 8, 9, and 10 (still to be given).    See *M. G. E.-V.*, pp. 402-3, 413, 422.    We may now perceive the solution of the Jevons-Laspeyres dispute.    It the two commodities are equally important over both the periods together, Jevons was right.    If they were as Laspeyres supposed them to be, Laspeyres was right.    But neither was fully right unless he expressly confined his solution to the particular supposition about the weights (or the quantities).    This was done by neither of them.    Cf. *M. G. E.-V.*, p. 354.

important commodities is only approximative even under the first two suppositions, and it can be no more than approximative for the present case.    The expectation that we could pass from the geometric mean to the geometric average has not been fulfilled.    Like Moses, we have just caught a glimpse of the promised land, and are then kept out of it.    But we have come near it, and we may abide near it. The geometric average *is* approximative in the present general case just as much, and in the same way—not cumulatively, but neutralisingly—as it was in those special cases.

Other approximations may also be made.    The geometric method is not the only one that can be taken over from the earlier suppositions to this third.    In the second supposition formula (6) can be used without the condition that $q_1p_1$ is = $q_2p_2$, that is, it can be used generally, or under our third supposition.[120]    And now it yields results, as tested by our circular test, as good, but no better, than it yielded there. This formula (6), being analysed as formula *b* above was analysed, may be shown to be (1) the arithmetic average of the price-variations with weighting according to the prices of the first period multiplied by the geometric mean between the quantities of the two periods ;  (2) the harmonic average of the price-variations with weighting according to the prices of the second period multiplied by the same geometric means. By analogy we see that it will also be approximately (3) the geometric average of the price-variations with weighting according to the geometric means between the full values of the two periods.    It is, then, approximately equal to formula (7). This is the inverse of what we knew before ;  for we have known that formula (7) is approximately equal to it.    In the second supposition formula (6) is better than formula (7). Therefore in general we may expect formula (6) to be better than formula (7)—better than the geometric method.

Another approximative method can be obtained by combining the best results from the two elementary suppositions, that is, combining formula (1), in which the quantities of the two periods must be distinguished, with formula (5), which cannot be taken over alone as could formula (6), into this :

$$\frac{P_2}{P_1} = \frac{q_2'p_2' + q_2''p_2'' + \cdots}{q_1'p_1' + q_1''p_1'' + \cdots} \cdot \frac{q_1'\sqrt{p_1'p_2'} + q_1''\sqrt{p_1''p_2''} + \cdots}{q_2'\sqrt{p_1'p_2'} + q_2''\sqrt{p_1''p_2''} + \cdots}$$

$$= \frac{\Sigma v_2}{\Sigma v_1} \cdot \frac{\Sigma q_1 \sqrt{p_1 p_2}}{\Sigma q_2 \sqrt{p_1 p_2}} ; \tag{8}$$

[120] Cf. *M. G. E.-V.*, p. 360.    This method is there called Scrope's

in which it is evident that if the quantities are the same in both the periods, it reduces to formula (1), and if the full values are the same in both the periods, it reduces to formula (5).    There is good reasoning leading up to this combination, too long to repeat here ; [121] and yet this formula does not satisfy the circular test any better than does either the properly weighted geometric average or formula (6), though it meets that test about as well as they do.[122]

Again : we have remarked that in the present extensive supposition the arithmetic average with the weighting of the first period and the harmonic average with the weighting of the second period yield different results.    The fact is, as trial shows, they yield results which, when tested by the circular test, are found to lie on opposite sides of the truth, and apparently equally above and below it proportionally.    This suggests taking the geometric mean between them, which works out as follows,

$$\frac{P_2}{P_1} = \sqrt{\frac{\Sigma q_1 p_2}{\Sigma q_1 p_1} \cdot \frac{\Sigma q_2 p_2}{\Sigma q_2 p_1}} = \sqrt{\frac{\Sigma q_1 p_2}{\Sigma q_2 p_1} \cdot \frac{\Sigma v_2}{\Sigma v_1}}. \tag{9}$$

This method still fails to satisfy the circular test ; but perhaps it satisfies it best of all.[123]    Note that it involves the arithmetic average, the harmonic average, the weightings of the first and second periods, and the geometric mean.    It is, too, neither more nor less than the geometric mean between two forms of formula (1) applied to the quantities of the two periods one at a time.    It seems to contain everything that could be desired.

Nos. 6, 7, 8 and 9, then, are four methods which seem to have much in their favour.    They all keep very close together, and each comes very near to satisfying the circular test ; and they err sometimes above and sometimes below the truth, so that their errors are not cumulative over a series of successive periods, but tend to neutralise one another.[124]

emended method, pp. 373, 543.    It would have been better if this method, in all its forms, had been called Lowe's method.

[121] See *M. G. E.-V.*, pp. 374–82.

[122] This method is analysed and compared with Drobisch's, Lehr's, and Nicholson's, in *M. G. E.-V.*, pp. 383–96, and with formulæ (6) and (7), pp. 396–407.

[123] This method is not formulated in *M. G. E.-V.* ; but in note 16, on p. 429, it is tested and found wanting, but it happens to meet this particular test better than do any of the preceding methods.    It was unduly overlooked in that work.

[124] For remarks on the comparative merits of three of them,

We have seen reason to believe formula (6) better than formula (7). Perhaps formula (9) is the best of the rest, but between it and Nos. 6 and 8 it would be difficult to decide with assurance. The circular test may provide the criterion. Only it will require many laborious operations before a decision can be made. Who will undertake the labour ?

It may be questioned whether any method can be found that will satisfy the circular test perfectly, at the same time showing other ear-marks of correctness ; for, as we have seen, the test can easily be satisfied by many wrong methods. The problem of axiometry may after all be mathematically insoluble with perfect precision, like the problem of squaring the circle.[125] Mathematicians can work out the value of the symbol $\pi$ to any number of decimals ; but practical workers rarely employ more than four or five. Surveyors, in working around a territory and coming back to their base, never strike it with perfect accuracy ; and even after adjusting their measurements they are not sure of their results to every millimetre. If, then, our best methods, after proceeding over many periods, when worked back to their starting point, do not fit into it with perfect exactness, we need not be discouraged. Their precision is sufficient for practical purposes.

Indeed, their precision is greater than is needed, when we consider the lack of precision that exists in the data that are presented to these methods to operate upon. Therefore even less precise methods may be adopted in practice, if more convenient and yet fairly agreeing with these theoretically best methods.

see *M. G. E.-V.*, p. 407. Unfortunately the fourth, formula (9), was not there discussed.

[125] Cf. *M. G. E.-V.*, p. 408. There in a note some hints are offered for possible improvement of formula (7). Perhaps the reason why that formula fails is because in the compound unit constructed that is to have the same exchange-value over both the periods together, there is error due to the fact that the full values at the two periods are reckoned in moneys of different exchange-value, and it is impossible to allow for that fact till the measurement of the variation of the exchange-value of money is completed, whereas that compound unit is needed for the making of that very measurement. Now, one of the exchange-values of money must be smaller than the other, and therefore its greater smallness should be allowed for. The arithmetic mean does not allow for it. The geometric mean allows for it too much. Therefore some intermediate mean is needed that will allow for it correctly.

All these four theoretically best methods contain a geometric feature, involving the extraction of square roots. These operations, even with the aid of logarithms, are laborious. Two methods, substituting an arithmetic for the geometric mean, have been proposed. The first, independently suggested by Professor Marshall and by Professor Edgeworth, has been actually recommended by the latter and by the British Association Committee.[126] The second has been merely suggested, likewise independently, by Drobisch and Sidgwick,[127] and has been used by Dr. Bowley, as will be seen presently. They are

$$\frac{P_2}{P_1} = \frac{\Sigma(q_1 + q_2)p_2}{\Sigma(q_1 + q_2)p_1},\tag{10}$$

and

$$\frac{P_2}{P_1} = \frac{1}{2}\left(\frac{\Sigma q_1 p_2}{\Sigma q_1 p_1} + \frac{\Sigma q_2 p_2}{\Sigma q_2 p_1}\right).\tag{11}$$

The first of these arithmeticises formula (6), and the second arithmeticises formula (9). Which is the better, theory does not say, although if (9) is better than (6), (11) is probably better than (10). To decide between them appeal must again be made to the laborious task of testing them many times by the circular test.[128] Sufficient trial of them has already been made to show that even in unfavourable examples they come pretty close to the theoretically best methods. This congruence, along with their practical convenience, is the proper, and the only proper, reason for recommending them ; and on this ground, formula (10), being easier, may be preferred to formula (11). Other methods, many of them in use to-day, are absolutely bad ard pernicious. In many of them the errors are so great

[126] First *Report*, pp. 249–50 ; Edgeworth's *Memorandum* thereto, p. 266. Cf. *M. G. E.-V.*, p. 543.

[127] Both suggestions were merely incidental. Sidgwick made his in a footnote, in which he added the remark that such a mean is wanting in " practical significance," *Political Economy*, p. 68 n. Likewise Wicksell has said that the use of this mean has only a " conventional meaning," and therefore he rejected it, *Geldzins und Güterpreise*, Jena, 1898, pp. 8–9.

[128] The single use of this test in *M. G. E.-V.*, pp. 427–8, shows a slight advantage in formula (11). Formula (10) is examined in that work at length, pp. 409–23. There it is shown that this method and Lehr's method apparently err almost equally on opposite sides of the truth, suggestive of drawing a mean between them, though this would be too laborious for practice.

and so cumulative, that they cannot be used in the proper way, in the " chain " system ; but must be measured from a common base, and then they give rise to all sorts of haphazard weighting, involving unknowable errors, when later periods (the more recent periods) are compared with one another.   All these should be thrown on the scrap heap.[129] The two before us are sufficiently easy ; and any one who will shirk the trouble of working them, should abandon the subject.   These two, then, may be declared unequivocally the best practical methods ; and it is only necessary for the practical workers to come together and choose between them.

Let us illustrate the matter by an example.   In the chapter on index-numbers in his *Elements of Statistics* Dr. Bowley uses an example with three commodities, one of which is made to vary rather violently in price and quantity, as follows :

| *First year.* | | s. | d. | *Second year.* | | | s. | d. |
|---|---|---|---|---|---|---|---|---|
| 6 quarterns bread at 6d. | | 3 | 0 | 7 quarterns at 5d. | . | | 2 | 11 |
| 4 lbs. meat at 7d. | . | 2 | 4 | 5 lbs. at 8d. | . | . | 3 | 4 |
| ½ lb. tea at 3s. | . . | 1 | 6 | 1½ lbs. at 1s. 4d. . | | . | 2 | 0 |
| | | 6 | 10 | | | | 8 | 3 |

On these data Dr. Bowley performs five operations, rejects three as resting " on no sound hypothesis," and retains the first two because " they may be regarded " as giving " inferior and superior limits of the index-number, which may be estimated as their arithmetic mean (80·5) as a first approximation " (pp. 226–7).   This final performance is in accord

---

[129] There is, however, one practical method which avoids those errors and the consequent condemnation.   This is the method used by Lowe and recommended by Scrope, and by many others since, notably by G. H. Knibbs, who calls it "the aggregate expenditure method," *Prices, Indexes, and Cost of Living in Australia*, 1912, pp. 11–14.   Professor Mitchell, who adopts the same name for it, describes it thus : " By this method the cost of an unvarying bill of goods is calculated at the varying prices prevailing during different years," *Index-Numbers and Wholesale Prices*, p. 160.   Its formula is No. 1 above, but with the $q$'s representing no real quantities, but only general estimates of the average quantities that have been used during certain years. Professor Mitchell recommends it because its results can be shifted from one base to another.   This is analogous to saying it satisfies the circular test.   But we have seen that this test does not prove the correctness of a method.   The truthfulness of a

with formula (11) (multiplied by 100 to bring the result into comparison with 100 as the starting point)—and perhaps Dr. Bowley is the first to put that formula into operation. But because of these " may be's " Dr. Bowley's readers can hardly be impressed with the validity of this result, and still less when they remember that he has belittled the use of weighting (and never once mentioned double weighting, which is here employed) except in cases of great variations and few items, such as this is, and for such cases has recommended the geometric average. Moreover, he himself casts doubt on the result by speaking of it as " a first approximation," though he gives no hint of any subsequent procedure.[130] We are not surprised, then, that in his latest utterance on the subject, in a paper on *The Measurement of Changes in the Cost of Living* in the *Journal of the Royal Statistical Society*, May, 1919 (pp. 348–9), Dr. Bowley says little in favour of this method (which there is given as Method III.), but thinks it " appears valid " for certain purposes. This is not the way to recommend a good method.

Now, on applying to these data the several theoretically best methods, worked out to the second decimal for distinctness, and giving Dr. Bowley's also in this form, along with the other practical method, we get the following results :

No. 6, giving index-numbers 80·92, indicating fall of 19·08 per cent.

| ,, 7 ,, | ,, | 81·07 | ,, | ,, | 18·93 | ,, |
| ,, 8 ,, | ,, | 80·80 | ,, | ,, | 19·20 | ,, |
| ,, 9 ,, | ,, | 80·32 | ,, | ,, | 19·68 | ,, |
| ,, 10 ,, | ,, | 79·34 | ,, | ,, | 20·66 | ,, |
| ,, 11 Dr. Bowley's ,, | | 80·47 | ,, | ,, | 19·53 | ,, |

If formula (9) be the best of the theoretically best methods, we see that Dr. Bowley's " first approximation," the practical formula (11), is better, here at least, than some of the theoretically best. Here again also it looks better than No. 10. Then let us recommend it firmly, and not with mere " may be's." [131]

To return to theory : would anything be gained by drawing an average of the results yielded by several methods ?

method should not be sacrificed to obtain this and some other merely apparent advantages. But in past epochs, when the annual variations of quantities cannot now be ascertained, this is by all odds the best practical method to use.

[130] Jevons likewise had spoken of his simple geometric average as " a first approximation," without making any second, *Investigations*, p. 122.

[131] *M. G. E.-V.* was published three months after the first

Hardly, as they have different merits.    All that we can do is to choose the best, after testing all the candidates ; for to average the others with the best, would only vitiate the result.    This is said in spite of Professor Edgeworth to the contrary, who has advocated such averaging of various methods in his first *Memorandum* (pp. 266, 297).    Only if two methods can be shown to err nearly equally on opposite sides of the truth, can a mean between them give a closer approach to the truth, and then this mean constitutes a new method, as in formulæ (9) and (11).[132]    Professor Edgeworth apparently conceives of all the methods as so many chance shots at a mark, and, having no knowledge of any way of testing them, he relies on the theory of probabilities that they are likely to err equally on opposite sides, so that their average is more likely to approximate to the truth.    But this procedure rests on the supposition that the observations are made with equal care, or else they must be weighted— and no astronomer or surveyor would include very carelessly made observations.    But the methods which Professor Edgeworth would average are not made with equal care : some of them have been invented without any care whatever, being mere first guesses.    They would, then, need to be weighted.    But nobody can reckon the comparative carefulness with which the different methods have been invented. Therefore the only way to compare methods is to try them on cases where the true result is known, whereby we can measure their comparative erroneousness and their comparative accuracy.    And then we should choose the best, and let the others go.

### PROBABILITIES AND THE MEDIAN.

Making reference, reported in the *Journal of the Royal Statistical Society*, May, 1919 (p. 366), to *The Measurement of General Exchange-Value*, Professor Edgeworth criticised that work for " attempting to dispense with the need of Probabilities."    If that work has any merit, it is because of that attempt.    That work aimed at precision, at the attainment of absolutely correct results as far as within our power— not at " the judicious compromise and happy ambiguity " of another writer, which Professor Edgeworth in a previous

edition of Dr. Bowley's work.    Consequently the means of testing the method he so slightingly recommended were in existence when Dr. Bowley revised his later editions.

[132] And compare what is said in note 128.

review of that work held up for imitation.[133]   Everything
in it about the theory of averaging price-variations is
demonstrative, except the calculations of the amounts of
error of methods proved not to be correct.   But for the
unexpected breakdown of the geometric average in not
fulfilling the promise of the geometric mean, that work would
have left no room for the use of probabilities in the theory
of the subject, but only in its application to practice.

   In that work the best methods of averaging price-varia-
tions were put to various tests—among them the circular
test, never before used for measuring their comparative
errors, and unfortunately never since made use of ; and they
were thus tested on hypothetical examples, and those
examples were mostly, as in the usual practice with other
writers, cases of two commodities varying in price (and often
in quantity) oppositely.   Two commodities were chosen
because they are the simplest, and because they are always
sufficient to disprove a method, though not always sufficient
to prove one.   They were chosen also because it is natural,
in balancing opposite forces or influences, first to use two,
and only later to extend the investigation to the case of
many.   And those two prices were generally supposed to
vary extravagantly.   This was in order that they should
serve, like Bacon's *instantiæ ostensivæ*, to magnify any error
that might lurk in the method under examination, and thus
lay it open to more ready detection ; and especially this was
done when the erroneousness of different methods was being
compared.   Nobody knows better than Professor Edgeworth
(for he has drawn attention to it [134]) that different kinds of
averages applied to the same many slightly divergent figures
yield nearly the same results.   Therefore little information
can be elicited from applying the different kinds of averages
to ordinary cases of price-variations ; and while we have
them in our power, as we have in hypotheses, it is the part
of wisdom to make them strongly divergent, that the averages
may be clearly distinguished.

   This is not Professor Edgeworth's method, and it seems to
have displeased him.   So much so that he has lost his
judicial balance ; for at the same time he made the above
criticism he added another, to the effect that of the work in
question " a great part was devoted to the investigation of a
formula proper to the case, in which there were only two

   [133] In the *Economic Journal*, September, 1901, p. 410.
   [134] In the *Economic Journal*, March, 1896, pp. 136–7.

prices ! " [135]   This is simply a misstatement of fact, all the
more astonishing as that work was the first to point out the
danger of relying on only two items in dealing with the
geometric average.[136]   But that work did make much use
of two prices of the sort mentioned, in negative tests mostly,
or else in comparative tests.   This is the point of offence to
Professor Edgeworth.   In the *Economic Journal* of March,
1896, Professor N. G. Pierson made an attack on index-
numbers, in which he used extravagant price-variations of
two groups of commodities described as " equally impor-
tant," their full values being posited as equal sometimes at
the first period, sometimes at the second, and sometimes over
both together.   Then he applied to these price-variations
the same system of index-numbers—the arithmetic average
with even weighting.   Naturally he got very discordant
results.   Thereupon, without making any attempt to solve
the difficulty, except by suggesting another worse system,
the error of which he discovered, he hastily concluded that
" all attempts to calculate and represent average movements
of prices, either by index-numbers or otherwise, ought to be
abandoned " (p. 131).   In the same number of that journal
Professor Edgeworth took up the cudgels in *A Defence of
Index-Numbers*.   How did he conduct the defence ?   Did he
point out that Professor Pierson, while using always the
simple arithmetic average, had really obtained the " equal
importance " of his groups in three different ways, only one
of which could be proper for that average, though the others
might be proper for the other averages ?   And did he show
that each of the three simple averages applied to the suppo-
sition fitted for it would yield the correct result, and that the
same would be given by the other averages (or by the geome-
tric at least approximately) when applied to the same
suppositions, each with its proper weighting ?   No, he knew
nothing of all this.   He perceived, indeed, and pointed out,
that different weightings were involved.   He put his finger
on the source of the trouble.   But he did not know the
remedy.   So he brushed aside the various suppositions,
calling them " odd " " in view of the sporadic dispersion [of
price-variations] which very generally prevails in this
world " (p. 133).   In their place he advised us to deal only

[135] This and the preceding are only repetitions of criticisms
made in the review already referred to, several pages of which
were devoted to the discussion of a problem in probabilities
entirely irrevelant to the main subject.

[136] Cf. above, note 68.

with " concrete, sporadically dispersed, price-ratios," mean-
ing many and not very divergent price-variations, as these
would show only " insignificant " discrepancies ; and to
make our experiments, not on " artificially simplified
examples," but " in the spirit of Probabilities," " *not
consciously selecting cases which will not work well* " (p. 135).
Such conscious selection of extravagant examples, simplified
by confinement to two articles, but with their discrepancies
magnified for better discrimination, we have seen purposely
made by Galileo in his controversy with Nozzolini ; and we
have seen Nozzolini throw them out as improbable absurdities
unfit for consideration by rational dialecticians, but really
because they were cases that would not work well according
to his view of the subject. It may be submitted that
Professor Edgeworth follows Nozzolini rather than Galileo
in his method of treating the subject. He would purposely
confine attention to examples which minimise discrepancies
and so hide the need of getting rid of them. If this is " the
spirit of Probabilities," it is not the spirit of science.

But Professor Edgeworth has two special reasons for
objecting to the use of only two prices. The one is that they
do not admit the use of the calculus of probabilities ; the
other, that they do not admit the use of another pet of his,
the median. Of the latter, later. As for the former, the
fact that an example of two price-variations excludes the
calculus of probabilities, merely proves that the calculus of
probabilities should not be invoked—until it is needed. But
Professor Edgeworth has made up his mind that " the
problem now before us, in its data, methods and result, is
germane to the Calculus of Probabilities." [137]   So convinced
is he of this that, as he tells us in another paper, if a case
should turn up in experience of two large groups of commo-
dities simultaneously rising and falling in price, he would
deny to the idea of a variation in the exchange-value of
money any determinate meaning. His words are : " That
in such a case the exchange-value of money has varied by so
much, would appear to me a somewhat indefinite proposition
—its subject deficient in logical clearness, and its predicate
in numerical precision." Think of it ! Because his demand
for " Probabilities " has no application to this case, he denies
to it significance ! He denies, too, that the problem in this
case admits of solution, for he adds : " On such a supposition

---

[137] *Ibid.*, p. 132. He asserts that this " is generally implied,"
but does not say by whom.

the objections which have been urged by a distinguished economist against index-numbers, that the results are widely different according as different species of averages are employed, would seem to me a fatal objection." [138]    That is, Professor Pierson's unweighted and trivial objections are to be allowed, if such things happen ! Most certainly Professor Edgeworth has not answered Professor Pierson—has not known how to answer him, and does not wish to know.    The balancing of opposite price-variations—how the rise of one price may compensate, or come short of or exceed in compensating, the fall of another—remains beyond his ken. Yet what about the doubly weighted method indicated by formula (10), which Professor Edgeworth has joined others in recommending ?  That method is perfectly applicable to this case, provided data about the quantities be also supplied. and it would give a result insignificantly discrepant from the results given by the theoretically best methods.    Would Professor Edgeworth throw it overboard also, merely because some other wrongly weighted averages might give widely different results ?   Again, Professor Edgeworth is one of those who hold that formula (1) is correct on the supposition that the quantities have not varied.[139]    Would he say that this is no longer true, or that it has no meaning, if only two prices vary ?  And how about Dr. Bowley's example, above cited ?  That has three commodities, to be sure ;  but when does the number used begin to have meaning ?  In that example the price-variations are, at least, not sporadic, so that Professor Edgeworth would still have to discard it. Professor Edgeworth would have done better to confine himself, as before, to denying the probability of such things happening.[140]    But he has admitted that it can happen by taking it into consideration ;  and now it seems strange if he cannot understand a problem which so many other economists—Jevons, Laspeyres, Messedaglia, etc.—have worked at, though unsuccessfully.  Just because another has succeeded where they have failed and he himself has not tried—is that a reason for him to continue looking at the subject through a mist ?  The assertion may be added : if the case he supposed should happen, and if in addition it

[138] In the *Economic Journal*, September, 1901, p. 408. Page 410 he says the widely different results may amount " as likely as not to 25 per cent."

[139] *Memorandum* to the first Report, pp. 264, 272, 293.

[140] This, in fact, he does again do, a little further on in the same paper, *Economic Journal*, September, 1901, p. 410.

happened that the two groups should change in price in inverse ratio, and if they were of equal importance over both the periods together, then, no matter how many other commodities remained unchanged in price, the exchange-value of money has remained constant. And this proposition has been proved and demonstrated, and is as certain (to repeat), and (we may add) as plain, as anything in Euclid. If Professor Edgeworth can disprove that proposition, let him do so. If he cannot, what value has his statement that he cannot see the meaning in it ? [141]

But here a concession may be made. Even if an absolutely correct method of averaging price-variations were discovered, or one as correct as the use of the arithmetic average in reducing observations (and have we not reached one as good as that ?), yet the application of it to practice would give occasion for error, just as even the application of an absolutely exact yard stick to measuring the length of a line would give rise to error. Now, *The Measurement of General Exchange-Value* might have been a little more explicit by making some allusion to these errors of appli-

---

[141] Professor Edgeworth seems to desire to retain what Mill called " the necessary indefiniteness of the idea of general exchange value," and says " we will never reach an exactly true method [of measuring it] . . . until we are able to handle and weigh final utility, or . . . ' esteem value,' as we do material commodities " (*ibid.*, p. 409). On the contrary, we shall never reach an exactly true method even of understanding exchange-value until we cease to mix with it the idea of esteem-value. Any one who accepts Mill's denial of the possibility of measuring general exchange-value (a denial due to ignoring the balancing of opposite variations) ought, like Mill himself, to refrain from working over price index-numbers. Professor Edgeworth further agrees with Professor Marshall that, in the world as it is, " an absolutely perfect standard [of exchange value] is ' unthinkable ' " (p. 410). There is, of course, some inconsistency in saying that something one is thinking about is " unthinkable " ; but the meaning seems to be that it is unthinkable, that is, it cannot be rightly thought that such a thing exists or can exist—that is, again, it must be thought or believed that the thing does not or cannot exist. But such a statement needs proof. Well, *M. G. E.-V.* has gone perhaps the furthest of all in showing that an absolutely perfect standard of exchange-value, in the complex economic world, does not exist, just as (see above, p. 103) there is no absolutely perfect value of $\pi$, or just as an absolutely perfect value of $\pi$ is " unthinkable," since we simply cannot think what it is. See in that work p. 408.

cation according to the theory of probabilities. It might have expatiated more on the inaccuracy of observations or measurements, and on the probability of the increase or decrease of errors. The theory of probabilities shows that if in measuring a given magnitude four times as many observations be used as at first, the accuracy of the result, as indicated by its " probable error," is doubled, that is, it increases with the square root of the number of observations. Again, although a surveyor in measuring 100 miles, involving 100 times more partial measurements, probably commits 100 times more errors and altogether 100 times greater amount (or arithmetic sum) of errors, than in measuring a mile, yet, because his errors probably fall pretty equally on opposite sides of the truth, some of his partial measurements being too great and some too small, and because in the greater length greater opportunity is offered for the compensation of opposite errors, the probability is found to be that in measuring the longer line the balance of his errors, his final aggregate error (the *algebraic* sum of all the errors) is only ten times greater, or relatively to the lengths measured it is ten times less, or again, in other words, his accuracy is ten times greater, than in measuring the shorter line ; that is, in general, the absolute error increases, the relative error decreases, and the accuracy increases, with the square root of the increase in the magnitude measured.[142] Now, these principles may be taken over bodily into the mensuration of price-variations, with the proper *mutanda mutata* of course, the numbers of commodities being substituted for the numbers of observations, and the relative numbers of periods for the relative lengths of the lines. Here it will appear that the probability is (1) that, even if we employ a perfectly correct method, the final errors which we shall inevitably commit in practice, being by the nature of the case relative, will decrease, and our accuracy increase, with the square root of the increase in the numbers of the commodities operated on ; and (2) that as the measurements advance over a course of years, each being compared with the preceding in a new measurement, and the whole being strung out in one line, the errors to which even the perfectly correct method is exposed in practice will increase from the starting period (unless

[142] Cf. Professor Edgeworth, in the *Journal of the Royal Statistical Society*, June and September, 1888, pp. 365, 367, 602, 627.

adjusted by direct comparison with it) with the square root of the number of years traversed.   The first of these propositions may be extended to all methods ; the second to some, but not to all.   The second can be extended only to the methods of which the errors are shown by trial to fall pretty equally on both sides of the truth.   It cannot be applied to those of which the errors fall always or mostly on one side ; for when such methods are used in the " chain " system just referred to, their errors will grow in almost geometrical progression with the arithmetical advance in periods.[143] That work omitted to elaborate these precise propositions. They are here supplied.   Perhaps Professor Edgeworth will be satisfied with this filling up of the gap left by that omission.   It does not appear, however, that very much is gained.   Common sense, especially as it may rely on the study of probabilities made by others, would give the correct injunction to employ as many commodities in our averaging as is possible ; which injunction was given in that work (p. 77) on the authority of Professor Edgeworth himself. Common sense also teaches that compensating errors do not increase as do errors that do not compensate.   That work made considerable use of common-sense probability on the occasions when it was needed,[144] although it did not employ the calculus of probabilities.

And continuing to use such common-sense probability, we may argue in this wise.   The arithmetic average is no better than the harmonic, and the weighting of the first period is no better than the weighting of the second ; therefore let us use all these things together, and the result will have the greatest probability in its favour.   Now, precisely all this is done, when the quantities are the same in both periods, by the single formula (1), which we have found to be true on that supposition   And it might further be argued : the geometric average lies between the arithmetic and the harmonic, therefore it with the weighting of both the periods (likewise geometrically averaged) will most probably give the true result, and therefore agree with the preceding.   And this, too, we have found to be approximately true on the same condition, and exactly true if the condition is still

[143] More specifically, methods are found that augment or diminish both rises and falls.   If in a cycle prices first rise and then fall back to their former positions, such methods may be right at the end, but in the middle they will exaggerate or underrate the rise.

[144] Pp. 223 n., 321–4, 332, 341, 365, 394, 404, 406, 422, 423.

further narrowed to two commodities equally important over both the periods. Furthermore, if we drop the restriction to quantities the same at both periods, we may still use combinations involving all the averages and weightings mentioned, and these will then be the most probable. Such combining is done in formulæ (6), (7), (8) and (9), and with most perfect mixture in the last, which seems to stand the circular test best, and if it does, confirms the probability. And formalæ (10) and (11) are only slight departures from these, a more convenient average being substituted for a less convenient in parts of them ; which substitution, when the formulæ are applied, as they usually are, to a fairly large number of commodities varying not very greatly in prices and quantities, will produce little discrepancy in the results, so that they too have probability in their favour. One of these is the very formula recommended by Professor Edgeworth in the second section of his *Memorandum* to the first Report of the British Association Committee, as also by that Committee itself. Yet Professor Edgeworth did not make this argument from probability for that method—in fact neither he nor that Committee gave any argument in support of that method.[145] If he himself did not use probability for deriving and defending the method he recommended, why should he require others to do so ? Or if he had used it, would he have improved upon the arguments which have been given for it in the preceding section, and which were first given in *The Measurement of General Exchange-Value*, although Professor Edgeworth failed to see them there.[146]

But Professor Edgeworth has made use of the theory of probabilities, not in connection with that practical method, but in connection with the geometric average, and through it with the median. For Professor Edgeworth not only follows Laspeyres in advocating an arithmetic average, but, quite independently, he follows Jevons in advocating the geometric average—he follows both Galileo and Nozzolini. He advocates the geometric average for a special purpose

[145] The only thing like an argument appears in his *Memorandum* to the Third Report, p. 150, where he speaks of the " twin methods " using the quantities of the two periods separately, and adds, " there is no reason to think this method [one of the twins] would be less accurate than its converse "—the other twin ; which at least suggests combining them.

[146] Had Professor Edgeworth noticed what is written in favour of his own method in that work (p. 423), he would have perceived that that work strengthened his position.

different from the special purpose for which he advocated formula (10). The speciality of purposes—or of *quæsita*, as he calls them—is a peculiar feature in Professor Edgeworth's treatment of the subject before us. In one of his papers, making a reference to Dr. Venn's *Logic of Chance*, Professor Edgeworth, perhaps influenced also by a remark of Jevons, writes : " The answer to the question what is the *Mean* of a given set of magnitudes cannot in general be found, unless there is given also the object for the sake of which a mean value is required." [147]   In another he again writes : " It is with the index-numbers [of prices] as with conduct ; in order to form a just judgment, we must always look to the under-lying idea and purpose." [148]   Apparently Professor Edgeworth views these two statements as equivalent. They are very different. The former is true. The latter is not true. The truth of the former has already been explained. The error of the latter consists in its putting a variety of purposes within a purpose. In averaging price-variations the purpose or object is given : it is to measure variations in the exchange-value or purchasing power of money. This is precisely the purpose which Professor Edgeworth declares to be the object of his *Memorandum* to the first Report of the British Association Committee ; which Committee itself had been appointed " for the purpose of investigating the best methods of ascertaining and measuring variations in the value of the

---

[147] *New Methods*, in the *Journal of the Royal Statistical Society*, June, 1888, p. 347. Jevons had previously in the *Journal of the Royal Statistical Society*, June, 1865, written : " It is probable that each of these [the three averages] is right for its own pur-poses, when these are more clearly understood in theory," *Investigations*, p. 121. As it stands, Jevons's statement is per-fectly true—and we have seen different purposes, or uses, for which each of these averages is proper. Jevons, however, seems to have had the present subject only in mind. Even so, if we substitute " on certain conditions appropriate to each " for " for its own purposes," his statement is true : see *M. G. E.-V.*, p. 355, and cf. above, note 119.

[148] *Recent Writings on Index-Numbers*, *Economic Journal*, March, 1894, p. 163. It was perhaps Professor Edgeworth who inserted a similar statement near the opening of the first Report of the British Association Committee, making them say that the theoretical part of their work was " to distinguish analytically the different purposes which may be subserved by constructing a measure of a change in the value of money, and then to show what formula, what particular mode of combining the statistical data, is appropriate to each case."

monetary standard." [149]    " There are many methods—not one method—of measuring and ascertaining variations in the value of money," says Professor Edgeworth in the same *Memorandum* (p. 259).    There may be many methods, but only one true or best method—the one which comes nearest to giving the true variation of the exchange-value of money. To get at this true variation in the exchange-value of money is the sole object or purpose of averaging the price-variations, just as to get at the true length of a line by averaging the observations of it is the sole object or purpose of averaging those observations.    There is no other object or purpose— that is, no other main or final purpose. [150]    The single object or purpose has been stated.    It is specific. [151]    No further varietal subdivision of it is needed, nor is it possible, so far as the method of averaging the price-variations is concerned, including the method of weighting them.

The only differences that can come into our subject are differences as to the money itself of which the variations in exchange-value are being measured. [152]    Is it the money of England or the money of France ?  Is it the money of a whole country or the money which passes through the pockets of one class of its population ?  Obviously different data must be collected in each of these cases—different quantities always,

---

[149] So runs the official title of that Committee.    Professor Edgeworth opens his *Memorandum* thus :  " The object of this paper is to define the meaning, and measure the magnitude, of variations in the value of money. "

[150] There may be some subsidiary purpose sometimes in averaging price-variations, as to study their dispersion.    But this is only in aid of the metrical purpose, and cannot be put on a level with it or be substituted for it.    It is conducted on the same price-variations and by the use of the same method of averaging. There may be such a subsidiary purpose in averaging observations of length, etc. ;  but nobody has supposed that this calls for a different method of measuring length.

[151] Dr. Venn has thus repeated his doctrine :  an average, being a fictitious value, cannot take the place of actual values " for purposes in general, but only for this or that specific purpose," *On the Nature and Uses of Averages*, in the *Journal of the Royal Statistical Society*, September, 1891, p. 430, again p. 432.

[152] There are other kinds of value, such as the esteem-value and the cost-value, of money, that may be measured ; and naturally the methods of measuring them, whatever they be, are different from the method of measuring the exchange-value of money.    But these are not in question.    Who cannot distinguish between these, ought not to dabble with economics.

different prices sometimes.   But the *method* of averaging the
data is the same in all of the cases.   We do not require
different methods for averaging prices in England and in
France ; and the method of averaging the prices of family
budgets should be the same as the method of averaging the
prices of the national budget.   Only the *collection of the data*
is different in these different cases, different being the quan-
tities and sometimes the prices that go into the budgets of
employer families from those which go into the budgets of
employee families, different those which make up the national
budget of England from those which make up the national
budget of France.   Of course for *political* economists the
exchange-value of the money of their country—of the whole
country, not of any one class in it—what Professor Mitchell
calls " the general-purpose index-number " [153]—is the prin-
cipal object of interest.   " Let debtors and creditors," well
says Professor Edgeworth, " regulate their private affairs by
a special index-number, if they like.   That is not the affair
of statesmen and financiers.   But the currency *is* within the
province of government." [154]   In the collection of the
commodities and of their quantities and prices many varieties
of detail have to be considered (see above, pp. 88–9)—and
these may be spoken of as different methods, if you care to
do so ; but they are different methods of collecting the data.
Totally distinct is the method of averaging the price-varia-
tions of these data once collected.   This is the method under
consideration.   Of this method there is no subdivision.   It
is simply the method proper for averaging what is presented
to it, for the one common purpose of measuring variation in
exchange-value of the money brought into relation with the
data.[154a]

[153] *Index-Numbers and Wholesale Prices*, p. 26.   Professor
Mitchell complains that we " cannot conceivably devise a single
series that will serve all uses equally well "—that is, all par-
ticular or private uses.   There is nothing peculiar about this in
political economy.   Every tax, every tariff, affects individuals
and classes differently.   The political purpose is to look after the
general interest.

[154] *Memorandum* to the Third Report, p. 135.

[154a] In fact the *measuring* proper is done in collecting the prices
and quantities at the different periods ; which provides the
particular price-variations and their weights.   These are the
analogues of the measurements made by surveyors and the
observations made by astronomers.   The inversions of the price-
variations provide the variations of the *particular* exchange-values

In no other subject of mensuration is there any differentiation of methods to be used according to different ulterior purposes held in view.   How, indeed, in *any* subject could there be two different equally correct methods of measuring any magnitude ?   If the different methods give different results, how could they both be correct ?   It does not help matters to say that the one is correct for one purpose and the other for another.   Should a civil engineer employ two different methods of measuring a line (or rather of reducing his measurements to a final result), and on obtaining different results should he say the one was correct for laying out a railway and the other was correct for laying out a highway, he would probably get few customers.   Of course different degrees of precision may be required for different purposes ; but this acknowledges that only the most precise method is the truest.   When the magnitude to be measured is the variation in the exchange-value of money (of the same money, naturally, at the same time and place, and for the same people), there is no more room for two different but equally correct results.   In other kinds of mensuration there may be variety of different instruments fashioned to fit different objects that are to be measured :  these different objects correspond to the different collections of data in our subject.   There may even be sometimes a secondary more convenient method of roughly making some measurements, always tested by the one true method, as when we measure elevations by the barometer :  such a method corresponds to the practical methods in our subject, substituted for the one best method, to which they should approximately conform.   Otherwise the purposes we have in view when we measure the length of something (to repeat this instance) have nothing to do with the method of measuring it :  what we are going to do subsequently with the measurements made does not enter into the making of them.

And so it is with the measurement of variations in the exchange-value of money.   To measure variation in the exchange-value of money is the object or purpose of averaging

of money.   The *averaging* of these data yields the final measurement, but is not a measurement itself, although it is of course an essential element in the final measurement.   In axiometry, where the averaging yields the inverse variation of the *general* exchange-value of money, the theoretical problem is concerned with the averaging.   And there can be two or more ways of averaging (as regards the best method) for different purposes or *quæsita*, no more in axiometry than in astronomy or in surveying.

price-variations. If there is any other object or purpose in averaging them, then that is something else, and does not belong to our subject. To seek within the operation of measuring variation in the exchange-value of money by averaging price-variations for an undefined number of different methods according to arbitrarily chosen ulterior objects or purposes called *quæsita*, is an illusory refinement.[155] In fact, a couple of Professor Edgeworth's eight *quæsita* involve merely varieties of ways of forming the collection of data,[156] and four others do not belong in our problem at all.[157]

[155] In a notice of *M. G. E.-V.* in the *Yale Review*, May, 1902, following Professor Edgeworth's lead, Professor Fisher criticised that work for seeking only one method, instead of taking different purposes into view ; and illustrated the matter by a very curious description of two methods of measuring the height of a car, when really different heights of its whole and of a part of it were measured. But in his own *Purchasing Power of Money*, published nine years later, in the sections devoted to the mensuration of that power, he, too, sought only one method of averaging (with another as substitute), and said that the elements which enter into the construction of index-numbers must be chosen according to the purpose for which the index-number is desired (pp. 204-5, 232 ; and p. 392) ; giving an instance of such a purpose, he said " the same formula " may be employed (as in the other case), but the terms of this formula " have different meanings," that is, apply to different prices and quantities. Professor Fisher's last position is the correct one. He also criticised that work for not using the theory of probability ; and in his own work he did not use it. He further criticised that work for devoting all its efforts to measuring variation in the exchange-value of money, instead of extending them to measure variation also in the esteem-value of money ; and he confined the treatment of mensuration in his own work to the measurement of variation in the purchasing power (= exchange-value) of money.

[156] The second differs from the first (" the consumption standard " of a whole country) by employing private budgets, and the eighth differs from it by employing all the quantities sold and resold in all the mercantile and financial transactions of a country.

[157] The third, fourth, and fifth are concerned with the questions whether the standard should itself vary with the national affluence, income, or capital ; and the sixth with the definition of the appreciation (then going on, now depreciation) which monetary reformers (then the bimetallists) wished to rectify by infusing more money into (now abstracting some from) the circulation. All these are questions which the author of

Only the first and the seventh require two distinct methods of averaging price-variations. These two are one too many.[158]

In his first *quæsitum* Professor Edgeworth recommends formula (10). In the seventh he recommends the geometric average. All the first six *quæsita* are expressly made without any hypothesis as to the causes of the changes in prices. The seventh drops this denial, and introduces a clause about the " changes [of prices ?] affecting the supply of money," while the eighth contains the seemingly opposite " hypothesis that a common cause has produced a general variation of prices." [159] Incidentally, we might have supposed that particular causes produce variations in the prices (the true prices, relatively to variations in the exchange-value of money [160]) of particular commodities, while a general cause with reference to commodities, but a particular cause with reference to money, is what produces the general variation in prices and the particular inverse variation in the general

*M. G. E.-V.* has discussed in a separate work, *The Fundamental Problem in Monetary Science.*

[158] Professor Edgeworth even has a third ; for in the *Economic Journal*, June, 1918, p. 197, he agrees with Professor Mitchell in recommending " the aggregate expenditure method " as " the best form " for the latter's " general purpose series." (Above, in note 129, that method has also been recommended for use in past times, as the best practical method feasible for such occasions. It is there not recommended for any different purpose, the single purpose always recognised being that of making the best measurement we can of variations in the exchange-value of money.) Exactly how Professor Edgeworth intended to relate this new " best " method with formula (10) already recommended as the best, is not stated. As it is only briefly mentioned in passing, we may ignore it.

[159] Apparently suggested by Jevons, who, for continuing to use the simple geometric average, gave three equally poor additional reasons, the third of which is : " It [the simple geometric average] seems likely to give in the most accurate manner such general change in prices as is due to a change on the part of gold. For any change in gold will affect all prices in an equal ratio," *Investigations*, p. 121. If by " change in gold " he meant change in the exchange-value of money, the last statement is true (cf. *ibid.*, p. 128). But if he meant change in the quantity of money (which is the way Professor Edgeworth seems to take it) the statement is wrong, as it ignores what Del Mar felicitously named " the precession of prices."

[160] Cf. *M. G. E.-V.*, pp. 466, 475.

exchange-value of money—the bias above referred to (p. 72) —*whenever* such variations occur.   However this be, causes that have produced a magnitude which we wish to measure, do not properly enter into the measurement of the magnitude. Professor Edgeworth recognises this by saying near the beginning of his first *Memorandum* that "the business of this Committee is to measure a fact, not to speculate about its causes or consequences" (p. 258).   Unfortunately he immediately abandons this restraint by restricting it ; for he subjoins "it is only in the simpler kinds of measurement that the metretic art can be entirely divorced from theory about its subject-matter."   And so now in the seventh *quæsitum*, though he had not done so in the first, he brings into consideration the quantity theory that, other things being equal, prices vary inversely with the quantity of money (p. 280).   This theory would seem to most of us to call especially for the consideration of the quantities of the commodities the prices of which are varying.   Not so to Professor Edgeworth.   To him it seems that the price-variations of particular commodities may be affected by the " other things," but the general price-variation of all commodities is a "residual phenomenon," [160a] affected by the quantity of money.   The general price-variation is obtained by an average of the particular price-variations, and according to Professor Edgeworth it is to be obtained from them alone, *simpliciter* and *per se*.   Thus he says : "it would be a significant operation "—significant, apparently, of the quantity of money (or of what ?)—" to take the average of all the price-variations, irrespective of the quantities of the corresponding commodities" (p. 281).   This gets rid at least of the usual kind of weighting, and suggests either even weighting or some other kind.   Then, "the problem before us may be thus defined.   Given a number of observations consisting each of the ratio between the new price and the old price of an article, to find the mean of the observations."   Finally, "the problem as thus conceived belongs to that higher branch of the calculus of probabilities which may be called the doctrine of error" (p. 282).   Here are several questions— about the average, the weighting, the distinction from the first *quæsitum* and its formula, and the invocation of the higher probabilities.   It is difficult to disentangle these various things.   So let us get down at once to the last, and investigate its meaning and validity.

[160a] The *Journal of the Royal Statistical Society*, June, 1888, p. 353 ; *Economic Journal*, March, 1896, pp. 132–3.

Would it not have been better to reverse the last statement quoted, and say that the calculus of probabilities belongs to the economic problem—if we choose to apply it thereto ? If we choose, we can apply it to anthropological measurements, as Quetelet did.   But Quetelet had his single method of making his measurements before he subjected—not his method, which he never did, but—his measurements to the calculus of probabilities.   And he did this only because he found an analogy, before unsuspected, between his measurements and the chances of black and white balls coming out of an urn in certain combinations.   Now, Professor Edgeworth, having made a collection of " price-observations " (which phrase means what is generally meant, not by " prices," but by " price-variations," [161]), arranged them in order, as Quetelet had done with his chest measurements of Scottish soldiers, and found that—they did *not* follow " the typical law of error " (of observations), but fell into an unsymmetrical, though regular, curve, as he had suspected they would because of their similarity with estimates.[162] He therefore concluded that they should be averaged geometrically, like estimates ;  and as estimates are averaged simply, or with even weighting, and because of some misapplied analogy with gravity and the motion of the solar system through the sidereal system, he inferred that price-variations ought also to be averaged simply, " after the manner of Jevons " (pp. 281–2, 288).

All this is very pretty.   The analogy with estimates sets up a rival method, by the evenly weighted geometric average, over against the unevenly weighted formula (10), which he has already recommended.   We might expect, then, that Professor Edgeworth would examine and compare them, to see whether the geometric average, with proper weighting, would agree with the other method, and if not, how much they would differ ;  and then weigh them in the balance and determine which is the better.   No, he adopts them both ! Only the one is for one *quæsitum*, and the other is for another ! And meanwhile what is the calculus of probabilities doing ? As regards the method of measuring variations in the exchange-value of money by averaging price-variations, it is

[161] The phrase occurs again in the second *Memorandum*, p. 206. The price-variations, however, have probably been put in the form of price-deviations from a common starting point, which is omitted, leaving only the deviated prices to be operated on.

[162] Here, in the first *Memorandum*, pp. 282–3, occur the passages quoted above, pp. 79–80.

doing nothing.    It had nothing to do with Quetelet's method of averaging chest measurements.    It has nothing to do with Professor Edgeworth's method of averaging price-variations. Professor Edgeworth tries his best to bring this calculus into the inquiry about the axiometric problem, but without success except to confuse the subject.    All that this calculus can properly do is to come in *after* the method has been adopted and the measurements made.    Then it may be applied to the dispersion of the price-variations around their average.    And if Professor Edgeworth would confine it to this use, there could be no objection to his doing so.    He is at liberty to experiment all he pleases with the dispersion of price-variations, to investigate their modes and medians and quartiles, etc.    Perhaps some interesting developments may come from this application of the calculus,—but we still wait for them.    In the hands of Quetelet some interesting results did materialise.    In the hands of Professor Edgeworth none have appeared.[163]

And the reason is clear.    Professor Edgeworth has mistaken the analogy.    Because Quetelet got what he considered an objective fact from his measurements—his observations, not on the stars, but on men,—which he called a type, and conceived of as having objective existence (which others have since controverted), Professor Edgeworth thinks that in averaging " price-observations " in the simple geometric way he is getting something more than he got by the complex formula (10).    Formula (10) yields only an " average," not a " mean " (in the Quetelet-Herschel-Jevons phraseology),— only an ideal thing, a mere general average variation of prices (although its reciprocal is a variation in the exchange-value of money, which might seem objective enough) ; but

[163] He makes a suggestion, however, and leaves it at that. Quetelet, we have seen, noticed that if a population were made up of two races, the dispersion of their height measurements (and of others too) might show two modes (which would make the curve representing such dispersion look like the silhouette of a Bactrian camel's back).    And we have seen that Bertillon got some positive result from this study of the dispersion of stature in the Department of Doubs.    This question has also exercised Professor Edgeworth.    Perhaps price-variations are not of one type, but of two types—perhaps their curve is not " monocephalous," but " bicephalous " (not one-humped, but two-humped).    " This is a question of fact," which he leaves for others to decide.    *New Methods*, in the *Journal of the Royal Statistical Society*, June, 1888, p. 368.

the simple geometric formula, because it conforms to the nature of price-variations as dispersing themselves geometrically, reveals, he thinks, a real thing, a cause. This latter operation he condescendingly calls "secondary" to the other as "the primary." But he conceives it as a close second. Its "position of high collateral dignity," he says, "is all the more deserved in that the secondary measure enjoys an objective or external character, which cannot, according to my view of the subject, be accorded to the primary *quæsitum*." [164] That it is viewed as yielding something objective is expressed also in his speaking of it as serving the "purpose of ascertaining a type." [165] More fully he once refers to "that variety of index-number which purports to determine a real quantity, a cause or characteristic, such as 'scarcity of gold,' in some more objective sense than a mere fall of prices on an average. The *quæsitum* in this case may be likened to a physical quantity which is to be ascertained from a set of measurements. The method accordingly presents certain peculiarities derived from the theory of errors-of-observation." [166] In other words, Professor Edgeworth thinks that in averaging (geometrically) not prices, but price-variations, we are performing an operation analogous to the operation of averaging observations. The latter operation is conducted for the purpose of getting the true magnitude—"a real thing." Therefore the former also aims to get a true and real magnitude—which we may grant, seeing that it is the true magnitude of the variation of the price-level and inversely of the exchange-value of money which is aimed at (at least when we perform the averaging properly, or fairly well, as by formula (10) ), although this is not his meaning ; for his meaning refers to some other "real thing," discoverable only by his geometric average—a type or a cause, never yet described in clear lineaments, but a something which he once adumbrated as hovering "between genuine being and not being," "a muffled or masked representation of some real entity" [167]— for such nebulous phrases are his delight. Now, this analogy

---

[164] In the *Economic Journal*, September, 1901, p. 415, cf. pp. 409, 413.

[165] In the *Journal of the Royal Statistical Society*, June, 1888, p. 360, again p. 363 ; and pp. 352, 367, the average is said to represent "a real thing," an "objective thing."

[166] In the *Economic Journal*, March, 1894, p. 163.

[167] In the *Journal of the Royal Statistical Society*, June, 1888, p. 352.

is false. We have already seen the true state of the case (above, p. 71). In averaging price-variations for the purpose of ascertaining and measuring variation in the exchange-value of money, we are doing something slightly analogous to what astronomers and surveyors do when they average their errors for the purpose of ascertaining and measuring the " probable error " of their observations or measurements. This " probable error " is far from being an objective reality, and certainly it is not a type or cause of anything, though it may be an effect of something (of our carelessness or imperfection). Or, better still, it has considerable resemblance to the operation of getting the centre of shots deflected from the bull's-eye by a common cause, the wind ; which new centre never was, and never can be, aimed at, and is not a real thing existing anywhere " in heaven or earth." Hence the failure of Professor Edgeworth's search after an objective reality in the sense of an independently existing entity. He has missed Quetelet's trail. And what good, now, the appeal to the theory of probabilities (which he so identifies with the theory of averages as to call it by the latter's name [168]) can do in our problem of axiometry, still remains to be seen ; and until the good it serves is shown, we may be excused from bothering ourselves with it.

Possibly, however, an objective thing may be disclosed by the geometric average of price-variations, even though the analogy is false which is offered as a base for it. In fact, it may be repeated that an objective thing—a variation in the general exchange-value of money—is very approximately shown by the inverse of the geometric average of price-variations when properly weighted (and perhaps equally well by formula (10) ). But Professor Edgeworth is now searching for some other objective thing, to be disclosed by the geometric average not weighted in the way we have seen to be its proper weighting. That other objective thing, conceived as existing apart from the price-variations, he has generally described very vaguely, but sometimes he has given it a body and a name, referring to it as " a scarcity of gold," which is only a particular way of describing the proper quantity of money " in relation to the work it has to do " ; [169]

[168] In the *Economic Journal*, September, 1901, p. 412.

[169] *New Methods*, in the *Journal of the Royal Statistical Society*, June, 1888, p. 355. Obviously scarcity of money may exist independently of the averaging of prices, just as the height of a star above the horizon exists independently of the astronomer's averaging of his observations of it ; whereas general exchange-

for he was writing in a period of monetary appreciation.   If it is anything so definite as this, we can only wonder why Professor Edgeworth on other occasions made so much of a mystery of it.[170]   Yet if we pin him down to this definite description of the " real thing " to be disclosed by his geometric average, it becomes merely a question of fact whether his geometric average does disclose such a thing. This question must be simply negatived.   Here of course is where the quantity theory comes into play ; for if prices vary inversely according to the quantity of money, we have only to measure the variation of prices to determine the inverse variation in the quantity of money.   But Professor Edgeworth knows that the quantity theory rests on the condition of " other things being the same."   He therefore tells us that a " prior operation " is necessary to " reduce the problem to the simple case of general prices varying, other things being the same," which extraordinary operation he leaves " to the currency doctors." [171]   As he abandons to others the most difficult part of the job, he has not solved it. Moreover, in a similar connection he truly says that " variation in the effect indicates, but does not measure, variation in the cause." [172]   Therefore if scarcity of gold is the cause of falling prices, its amount is not measured by the amount of their fall.   Finally, if the purpose of averaging price-variations geometrically were to measure variation in the relative quantity of money, its purpose would be different from that of measuring variation in the exchange-value of money, and would not belong to axiometry.

value of money cannot exist independently of the particular exchange-values of money, of which it is composed, and which are inversely indicated by prices.   But this does not prevent the general exchange-value of money from being as objective as are its particular exchange-values.   See *M. G. E.-V.*, pp. 10, 12–13. Of course the general exchange-value of money that is measured is as general only as are the quantities of the commodities taken into consideration compared with the totality of commodities. So the general run of the stature of a race, rendered definite by an average, can be measured only from samples, as already remarked.

[170] Perhaps it is because the " scarcity of gold," being a cause rather than an effect, belongs as the object to the eighth *quæsitum*, leaving the " real thing " which is the object of the seventh *quæsitum* entirely undefined.

[171] *Loc. cit.*

[172] *Ibid.*, p. 352.

Still, the geometric average has been advanced as one belonging to the subject before us, and we must consider the question of the use in our problem of the geometric average either with even weighting or weighted as Professor Edgeworth would weight it. Here Professor Edgeworth has been misled, not so much by an analogy wrong from the beginning, as by one which he either does not follow or would still follow beyond the point where it ceases. The model he has set up for imitation is what should be done with estimates, which itself is modelled upon what is done with observations. Now, astronomers and surveyors average their single observations of the same magnitude with even weighting, provided they have made them with equal care. But they have two ways of weighting their observations of the same thing. The one is when two or more observations turn out alike : then they treat them as one observation and weight this with a coefficient corresponding to the number of the repetitions. As already quoted, weights are " merely numbers denoting repetition." The other is when two or more sets of observations of the same thing—the different sets made perhaps by different observers, or with different instruments—are themselves to be combined and averaged : then the procedure is the following. In each set its " probable error " is calculated, which shows inversely its precision. If the probable errors are equal, the sets (supposed to contain the same number of observations) are averaged with even weighting. If the probable errors are unequal, the sets are averaged with weights inversely proportional to the squares of their probable errors ; for when one set is, for instance, twice as precise as another, it is as if it contained four times as many observations made with the same care as those in the other, wherefore it must be weighted as four to the other as one.[173] The principle is the same in both cases. Let us take the latter model first. If Professor Edgeworth should desire, for instance, to combine two sets of measurements of the same money over the same two periods, say Mr. Sauerbeck's and *The Economist's* index-numbers, he might calculate, if he could, their probable errors, and weight them accordingly. As a fact, we rarely care to make any such combination of two sets of index-numbers. We generally prefer to accept the one we belive the better and let the other go. Remains the first method of weighting observations. Now, this model has been imitated above in the section

<hr>

[173] Cf. above, pp. 67, 113.

devoted to weighting, where we found that in our subject the individual can only be a certain quantity of value, wherefore commodities are to be weighted according to the repetitions of this quantity in their full values (with extra complication caused by the question of periods). And this kind of weighting has been employed in all the subsequent formulæ for averaging price-variations, including formula (10), as recommended by Professor Edgeworth, though it may be questioned whether he understood the full nature of that formula. Professor Edgeworth, however, in the present part of his work, takes for a single or individual price-variation, corresponding to a single observation of an astronomer or surveyor, the price-variation formed by the price-quotations at two periods of every single commodity that is entered in his list. But a single commodity is only a name for a greater or less number of objects, or of money-unit's worths of objects, so that it is sheer nominalism to treat every name of a commodity as an individual thing. Yet in his use of the geometric average in imitation of the use of the arithmetic average by Herschel or by Quetelet, Professor Edgeworth would give no more weight to the price-variations of the multitudinous money-unit's worths of wheat than to the scanty money-unit's worths of cloves. How far he wanders from true imitation of his model, should be evident. And it should be evident that he has no right to depart from it in the formula for his seventh *quæsitum* any more than in the formula for his first *quæsitum*, where he kept to the model. Or rather, he has no right to lay down a *quæsitum* " irrespective of the quantities of commodities," since this has no rational meaning.

But Professor Edgeworth is not altogether satisfied with even weighting. The oft-repeated condemnation of weighting wheat and cloves alike was perhaps too much for him. He makes concessions, and speaks of allowing some different weights to the price-quotations. Only they must still be different from the weights used in formula (10) ; they must not be regulated by the importance of the commodities to the consumers (or to the nation)—not by the numbers of money-unit's worths repeated in them. Why this continual refusal to weight the geometric average as he weighted formula (10) ? Because astronomers and surveyors allow two similar observations to count as two only if they are independent of each other ; but a double quantity of wheat (over against barley) has similar prices not independent of each other, since the same commodity always commands the

same price in the same market.[174]    Here is where Professor Edgeworth clings to an analogy no longer existing.    For as a matter of fact the price of barley is also not independent of that of wheat.    And in wheat itself a factor of its price is its material quantity.    A new element here comes in, which has no existence in the problem of averaging observations. This new element differentiates our present problem from that problem, and henceforth the analogy ceases ; for to leave the new element out is simply to shut one's eyes to a factor which exists all the same.    But there is one principle which underlies all averaging, which is that the individuals averaged must be weighted according to their repetitions. Now, at least the two quantities which make up the double quantity of wheat are independent of each other, and to that extent the full values which make up the double full value of wheat (compared with that of barley) are independent of each other.    This is as far as the alleged total independence of observations can be carried.    And there is no need of carrying it further.    There is a good reason why observations must be independent ; for their individuality is conditioned by their independence.    But in prices this is not so : the individuality of the price-variation of one quantity of wheat exists notwithstanding that the price-variation itself must go hand in hand with the price-variation of some other quantity of wheat.    Thus independence of the items, though necessary in observations, is not necessary in prices.    We average many things that are not totally independent.    The heights of men of the same race, for instance, are not totally independent, since they are dependent upon the physical conditions to which their ancestors have been subjected. For that matter, the observations of astronomers and surveyors are not totally independent, since they depend upon three common factors—the object itself, the instruments, and the observer.    They are independent only as separately made. Independence as well as individuality differ in different objects.    But individuality, not independence, is the universal element.    Professor Edgeworth has inverted the true relation, giving to independence the precedence which is due to individuality.

But having rejected weighting according to full values or importance, and being not altogether satisfied with even weighting, Professor Edgeworth looks around for another system of weighting, and finds it in the dispersion of the

[174] Cf. *Memorandum* to the First Report, p. 287.

prices. The quantities of the commodities are still to be excluded : the weighting is still to be irrespective of them. Thus in his first *Memorandum* (pp. 286, 287–8) he speaks of weighting the price-variations inversely according to the width of their dispersion, that is, according to their own amounts, so that commodities whose prices fluctuate little shall have greater weight than those whose prices fluctuate much. Herein he adopts one of the careless suggestions thrown out by Jevons (see above, p. 87). The width of a commodity's price-variation is considered a gauge of that commodity's importance, not to its consumers, but to " the calculators of probabilities " ! [175] This is said to be on the principle that according to the theory of errors of observation (and of estimates) " in the combination of the given observations less weight should be attached to observations belonging to a class which are subject to a wider deviation from the mean." [176] But that rule of astronomers and surveyors refers to the combination of different sets of observations of the same magnitude, not to the averaging of different observations in the same set, and so is an entirely different affair, misused as a model here. Perhaps, however, this is not his meaning ; for then his weighting would have had to be inversely according to the *squares* of the price-variations, treated as " probable errors." Perhaps he, and Jevons before him, had reference to a problem of true errors like that which we have discussed above, and in which we found, in effect, that the weighting had to be inversely according to the relative amounts of the errors (p. 62) ; which may here be assimilated to the relative amounts of the fluctuations of the prices. But that, too, was another considerably different problem. Price-variations, put in the artificial form of price-deviations, have some resemblance to errors ; but they are not errors, and have differences too. And these differences are especially influential in this very subject of weighting, as we have already seen (above, pp. 81–2). The mere fact that a certain weighting is proper in one problem, is not a sufficient reason for introducing it into another. In that problem of true errors we saw plainly the reason for giving a greater weight to the smaller error. Professor Edgeworth here says that " if more weight attaches to a change of price in one article rather than in another, it is . . . on account of its importance . . . as affording an

---

[175] *Memorandum* to the Third Report, p. 157.
[176] Repeated in the *Economic Journal*, June, 1918, p. 188.

observation which is peculiarly likely to be correct." [177]
This is hardly so.   The data of a commodity whose price
has varied 5 per cent. from one period to another are not
more likely to be correct than the data of a commodity
whose price has varied 10 per cent.   Certainly they are not
likely to be twice as correct in the former as in the latter.
The consideration of likelihood or probability brings the
analogy back to that of probable errors ; and now our
astronomers and surveyors do not attach more weight to
observations near the average than to observations further
from it, notwithstanding that the former are more likely to be
correct than the latter.   Professor Edgeworth both follows
and does not follow his models.   It is difficult to make out
his meaning with regard to his new kind of weighting ; and
it is needless to do so, as he reverts in the end to even
weighting.[178]   It is mostly with the median that he would
use his new kind of weighting.   To the median, then, let us
turn.

Professor Edgeworth has noticed in this connection what
he calls " the curious fact, that the median seems to keep
closer to the geometric than [to] the arithmetic " average.[179]
There is nothing curious about it (provided the data be
numerous enough), if the data are those which arrange
themselves according to the geometric law of dispersion.
The median then follows the geometric average, as it follows
the arithmetic average when the data arrange themselves
according to the arithmetic law of dispersion (cf. above,
p. 43).   Observations arrange themselves in the latter way,
estimates in the former ; and prices have been found to
arrange themselves in the former way too.   And now, for the
same special *quæsitum* for which he recommended the
geometric average, he recommends the median as following
it.[180]   He uses even an opposite reason for it, because of its
departure from the geometric average on some occasions,
when the latter is unduly influenced (and the arithmetic
average still more) by some abnormally high price-variations,
as of cotton in 1862, which troubled Jevons ; for then he
thinks the median, being unaffected by their extravagation,
would have given a better result.[181]   But we have seen that

[177] Third *Memorandum, loc. cit.*

[178] First *Memorandum*, p. 288.

[179] *Memorandum* to the Second Report, p. 208.

[180] *Ibid.*, pp. 208–9.

[181] *New Methods*, in the *Journal of the Royal Statistical Society*,
June, 1888, pp. 358, 360, 362.

that trouble was due, not to the geometric average in itself, but to its use with wrong weighting (above, p. 87). This, however, is a minor detail [182] for extraordinary occasions. For ordinary occasions the principal argument is the median's fellowship with the other average. There is even a double-headed argument for the median. For if it turns out after all that he was mistaken and price-variations do behave according to " the normal probability curve," that is, the arithmetic, then the median follows the arithmetic average ; and if he was right and price-variations behave according to the revised (logarithmetic) probability curve, the geometric, then the median follows the geometric average ; and so it is the safest, as it is close to being right in either case.[183] This is the argument of one who is not sure of his position. But perhaps his idea rather is, that prices do disperse themselves in both these ways on different occasions, and therefore the median is the best, as it automatically adjusts itself to either behaviour of the prices.[184] The median, then, is " *the* method "—" the most comprehensive combination," the best " compromise between all the modes and purposes." [185] Apparently it is recommended also to replace formula (10) : its following the arithmetic average as well as the geometric justifies its ousting them both.[186] The argument is lax, because the median follows the one or the other with any highly probable closeness only in cases when a great many items are taken into consideration. Otherwise it follows the others erratically, as Professor Mitchell, who has much dealt

[182] Professor Fisher, who likewise recommends the median on top of another method already recommended as the theoretically best, urges another detail in its favour : that " it has the advantage of easily exhibiting the tendency to dispersion of prices," *Purchasing Power of Money*, p. 231, again p. 427. If there is any need of such an exhibition, there is nothing to prevent any statistician from arranging price-variations in their order of magnitude and studying their dispersion, whatever be the method he adopts for measuring variation in the exchange-value of money. Why not keep distinct things separate ?

[183] Cf. *Memorandam* to the First Report, p. 291.

[184] Cf. *New Methods*, in the *Journal of the Royal Statistical Society*, June, 1888, p. 362.

[185] First *Memorandum*, p. 296.

[186] But we have seen in note 158 that Professor Edgeworth admits the median is not so good an average as another for at least one purpose not clearly defined. What his exact opinion is, cannot be fathomed.

with it, has found by experience, as might have been expected from theory, since " its precise position is often dependent on the relative price of a single commodity, which stands in the middle of a scale of relative prices." [187]   Yet Professor Edgeworth would allow its use applied to only forty-five commodities, that is, to forty-five price-variations.[188]   With such small collections, close agreement of the median with either the geometric or the arithmetic average can be expected only in the long run.   " But we must remember "—it is Professor Edgeworth himself who says this—" that what is true only in the long run is apt not to be serviceable on a particular occasion." [189]   What opinion, indeed, would people have of an astronomer or a surveyor who should use the median ?   " To take the median for the sake of avoiding computation," says an authority on the subject, " can only be justified when the observations are rough ones, and then the median itself is liable to differ considerably from the arithmetical mean.   The uncertainty of the probable error of the median is greater than that of the arithmetical mean, 217 observations being necessary in the former case to give the same [as little] uncertainty as 100 observations give in the latter case." [189a]   What is here said of the median in comparison with the arithmetic average in connection with observations, applies similarly, though with somewhat differing proportions, to the median in comparison with the geometric average in connection with estimates and with price-variations.   To be sure, in the field of price-variations (and still more of estimates) the data collected are only rough approximations themselves ; yet this is not a good reason for avoiding the trouble of computation, with risk of raising the uncertainty by nearly half.   Professor Edgeworth himself admits that the probable error of the median is slightly greater than that of the arithmetic average ; [190] and for measuring the general trend of prices he has maintained that the arithmetic average is less accurate than the geometric. Then why all this advocacy of an average admitted to be the least accurate ?

[187] Mitchell, *Gold Prices under the Greenback Standard*, 1908, p. 58, cf. p. 33 n. ;   cf. *Index-Numbers and Wholesale Prices*, pp. 85, 90.

[188] *Memorandum* to the Second Report, p. 209.

[189] *New Methods*, in the *Journal of the Royal Statistical Society*, June, 1888, p. 350 n.

[189a] Merriman, *Method of Least Squares*, § 167.

[190] In the *Economic Journal*, June, 1918, pp. 193-4.

But here again the consistency of his argument is broken by his allowing some uneven weighting to the median.   Or is it only when the median is supposed to take the place of formula (10) that uneven weighting is to be allowed ?   But then the automatic working of the median, which was supposed itself to choose its own guide, is interfered with. This does not seem to be his position.   But what the weighting is to be, it is difficult to make out.   It must not be according to the importance of the commodities, or to their relative " masses of value " ;   for, " from the present point of view, in order to determine a type by the Calculus of Probabilities, the attribute which gives importance to an article is not its quantity, but the independence of its fluctuation.   From this point of view pepper may be as good as cotton," [191] that is, apparently, cotton need have no more weight than pepper.   Yet " upon this ground," he " recommended assigning rather larger weights to the more massive commodities," that is, more to cotton than to pepper ! " For technical reasons I advise that each price-variation should count for as many, should in effect be repeated as often, as the square root of the number of units in the quantity of the corresponding commodity " (ibid.).   What the units are, is not specified.   Nor are the technical reasons given ;   but we may suspect one to be the familiar law that closeness of practice to theory, or the diminution of error, is increased proportionally to the square root of the number of items employed.   The idea may be that the quantities of the more massive commodities, being more numerous or more often repeated, yield more accurate results in this proportion.   But we have seen that they should then be weighted according to the squares of their probable errors. This would bring their weights back into direct proportion with their relative massiveness or importance or full values (for we cannot believe he referred to physical masses), though the question of the periods is left untouched.   Still, if weighting is to be allowed to the median, it might equally well be allowed to the geometric average ;   and if we used the weighting of both periods, we should be using formula (7), which Professor Edgeworth never dreamt of.   At all events, no superiority of the median over the geometric average would now appear : rather, the median would sink back into its proper inferior position as a mere more or less

[191] *New Methods*, in the *Journal of the Royal Statistical Society*, June, 1888, p. 363.

faithful attendant. The median can be advocated only because it is more convenient to consult than is the average it is supposed to attend. But a weighted median is again a difficult average to work with, and the margin of convenience is much reduced. As for the geometric average, this, with its proper weighting (as in formula (7) ), does not need to be confined to any such special *quæsitum* as Professor Edgeworth would confine it. Professor Edgeworth did not know that, applied to the data collected for his first *quæsitum*—the practical part of his subject, the case for which he recommended formula (10),—the geometric average with its proper weighting even in extravagant examples closely approximates to (and in practice so closely as to be almost indistinguishable from) that other formula. Had he known this, he would have had no occasion to distinguish the cases. Instead of following both Galileo and Nozzolini, both Jevons and Laspeyres, when they differed, he would have seen that in this our subject, because of the peculiarity here of the weighting, the positions of these opponents, when properly laid off, come so close together that their difference is negligible.[192]

One root of trouble in all this matter is the unwillingness of Professor Edgeworth and his followers to give due attention to weighting even in theory, because they think it of little importance in practice. Even in practice, however, it is important. Professor Edgeworth in his *Memorandum* to the Second Report of the British Association Committee studied the discrepancies of averages differently weighted, and came to the conclusion that " other things being the same, the inaccuracy of the price-returns affects the result more than inaccuracy of the weights " (pp. 197–8). As " a practical conclusion " he therefore advised : " Take more care about the prices than the weights " (p. 200). Yet this is not a good reason for not giving all the care that we can to weights ; and

---

[192] Professor Edgeworth not only did not know this and some other things when he invented his theory, but he continues in the same ignorance since the publication of *M. G. E.-V.*, which he reviewed and criticised and presumably read. So enamoured is he of his own probabilistic way of looking at the subject, that he refuses to be instructed, and shuts his eyes to the truth, just because it has come from another mode of treating the subject more analytical (dialectical he called it) and systematic than his own. Yet where plain algebra is sufficient for dealing with a problem, it is simple pedantry to lug in the so-called higher mathematics.

it is only perverted when Dr. Bowley says " Do not strain after exactness in weighting." [193]   A consequence has been that, although in their Reports the British Association Committee always recommended the use of uneven weighting, they did so half-heartedly, and damned it by adding that, notwithstanding their recommendation of it, it was " in one aspect [whatever that refers to] almost an unnecessary practice to secure accuracy." [194]   What is meant by " accuracy " here may be judged by the fact that in the annexed *Memorandum* their secretary, Professor Edgeworth, had characterised as small such discrepancies as those of 110·4 and 115, 94·6 and 92·4, 119 and 116, etc.[195]   If a government, whose business it *is* to regulate the currency, were to use index-numbers liable to such errors, which might accumulate from year to year (for the Committee rightly recommended what has since been called the " chain " system of making index-numbers for every year compared with its predecessor [196]), that government .would do better to leave the currency alone.   Evidently the statistical problem has been injuriously affected by such want of precision ; and few statisticians have paid attention to the recommendation of the Committee, despite the fact that  the formula they recommended is a very good one, and perhaps the best (unless formula (11) is better) for practical purposes.

Is it not time, then, that the Report  of the Committee should cease to be spoilt by the imperfect theoretical treat-

---

[193] *Elements of Statistics*, p. 118.   More recently he has written : " It does not seem at all certain that any improvement is made by weighting," in the *Journal of the Royal Statistical Society*, May, 1919, p. 346.   Compare with the above the practical advice given in *M. G. E.-V.*, p. 180 n.

[194] Second Report, p. 185.

[195] *Memorandum* to the Second Report, pp. 202, 206.   In the *Economic Journal*, June, 1918, p. 185, Professor Edgeworth disputes Professor Mitchell's statement that the  " chain " system is more trustworthy than the " fixed-base " system, by attempting to show on an example that the difference is slight. The one system indicates a rise from 111·5 to 113, and the other one from 126 to 130.   The latter is the same as a rise from 111·5 to 115.   The difference between 113 and 115, or between rises of 1·35 per cent. and 3·13 per cent., is treated by Professor Edgeworth as a " remarkable consilience . . . greater than was to be expected " !   Evidently he would have been content with an even greater discrepancy.

[196] First Report, pp. 250, 252.

ment with which it was accompanied ? Professor Edgeworth has won a high place of authority in this subject of averaging price-variations.    But it may be contended that the recognition of his authority has stood in the way of progress.    A five-fold charge may be brought against him.    He has led others astray (1) by underrating the importance of weighting as well in theory as in practice, (2) by complicating the subject with a variety of *quæsita* artificially constructed and apparently unlimited in number, (3) by ignoring the balancing of opposite price-variations, (4) by insisting on introducing the theory of probabilities in places where it does not belong, and (5) by winding up always with the median because of the mere probability that it follows close upon the average for which he would substitute it.

Would it not be well if statisticians and economists should again come together and decide authoritatively on the proper method of constructing index-numbers ?    And as regards the formula to be adopted, why should they choose any other than the one recommended by the Committee, unless indeed it be formula (11) ?    If this had been done a few years ago, it might have prevented such a regrettable relapse as that of Professor Fisher, who has tried to revive Paasche's use of the quantities of the second period, in a form which really is the harmonic average of the price-variations with the weighting of that period.[197]    It might also have prevented the still more regrettable, since practically misleading, relapse of the Labour Department in their *Labour Gazette*

---

[197] Cf. above, note 118 ; *The Purchasing Power of Money*, pp. 201–3, 218, 231–2, 421.    Professor Fisher used four tests (each with a duplicate), of which the first is incomplete, the second unnecessary, the third questionable, and the fourth is Westergaard's test.    All these tests can be satisfied by very bad methods. But the circular form of the last test—the most important of all, since it alone measures the amount of error in a method—he did not use.    He therefore never measured whether a method failed to meet his tests only slightly or very badly.    He apparently did not notice the criticism of Paasche's method in *M. G. E.-V.*, pp. 193–4, 541–2, nor the bad failure of that method to meet the circular test in an example on p. 428.    But in his recently published *Stabilising the Dollar*, he has dropped that method (and the median too), and recommended, pp. 3–4, 85–6, 206–7, Lowe's " tabular standard " method referred to above in note 129. This, really a multiple standard, though not the best method, is of course a great improvement on the single gold standard which it is intended to supplant.    But cf. p. 494 of the first work.

into revival of Laspeyres's equally faulty use of the quantities and weighting of the first period, with arithmetic averaging. The subject is one that is growing in importance. Labourers are more and more clamouring for regulation of their wages by index-numbers. The use of them has been recommended to bankers for the regulation of their issues. Professor Fisher himself is advocating their use for stabilising the monetary standard. The day for slip-shod methods, it would seem, is passing.

# INDEX